UNIVERSITY OF CALIFORNIA SAN DIEGO

Understanding Many-Body Effects in Ion Hydration with the MB-nrg Potential Energy Functions

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Chemistry

by

Marc Riera Riambau

Committee in charge:

Professor Francesco Paesani, Chair
Professor Michael Holst
Professor Andrew Kummel
Professor Stanley Opella
Professor John Weare

2019

The Dissertation of Marc Riera Riambau is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and most important, I thank my advisor, Professor Francesco Paesani. You have become a reference as a researcher and an ethical role model for me. Thank you for always being there for me, for all the knowledge you have help me to acquire and transmitted to me, for your guidance and support even on the hardest periods during these amazing years, and most importantly, for being such an amazing mentor. It has been an honor to be part of the group for all these years, and have been able to contribute actively and passively to develop and expand all the amazing science we have done in the group. And thank you for pushing me out of my comfort zone, helping me to stay there and teaching me how to make that new area mine.

I would also like to acknowledge Dr. Andreas Götz for his help and useful scientific discussions and advice, especially during the first years of my doctoral degree. Thank you for always being there for me, even in your busiest periods of time. I thank all the members of the Paesani lab, both past and present, for their support and help during these years, and very productive science discussions.

I also thank Dr. Daniel Smith and Dr. Andrew Simmonett. Their help and advice over these last two years has been key in developing my coding skills and good practices in software development. The MBX software would have never been possible without you.

I also thank CAICE, the Center for Aerosol Impacts on Chemistry of the Environment - An NSF Center for Chemical Innovation - for funding me the first years of my doctorate, and all the people involved in the center for their input in my research. And I also thank MolSSI, the The Molecular Sciences Software Institute, and all the software scientists involved, for their funding and mentorship in the software development part of the dissertation during the last two years of the doctorate.

Chapter 1, in full, is a reprint of the material as it appears in "i-TTM Model for Ab Initio-Based Ion–Water Interaction Potentials. II. Alkali metal ion – water potential energy functions", Riera, M.; Götz, A.W.; Paesani, F.; Phys. Chem. Chem. Phys., 18, 30334, 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in "Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces", Riera, M.; Mardirossian, N.; Bajaj, P.; Götz, A.W.; Paesani, F.; J. Chem. Phys., 147, 161715, 2017. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in "Isomeric Equilibria, Nuclear Quantum Effects, and Vibrational Spectra of $M^+(H_2O)_{n=1-3}$ Clusters, with M = Li, Na, K, Rb, and Cs, through Many-Body Representations", Riera, M.; Brown, S.E.; Paesani, F.; J. Phys. Chem. A, 122, 5811, 2018. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in "Many-Body Effects Determine the Local Hydration Structure of Cs^+ in Solution", Zhuang, D.; Riera, M.; Schenter, G.K.; Fulton, J.L; Paesani, F.;J. Phys. Chem. Lett., 10, 406, 2019. The dissertation author was the primary investigator and author of this paper.

Finally, I would also like to thank my fiancee Yazmin Tenorio, and my parents Antonio and Francina, for their moral support. I also thank all my siblings, specially Esther, for all the afternoons and nights I have hanged out with her talking about work and life, and all my friends who have been supporting me over this journey, with a special mention to Laura and Paloma, who have helped with enjoying life as I was also enjoying the science.

VITA

2012	Bachelor of Science in Chemistry, IQS School of Engineering, Catalonia
2014	Master of Science in Theoretical Chemistry and Computational Modeling, Autonomous University of Madrid, Spain
2014-17	Teaching Assistant, Department of Chemistry and Biochemistry, University of California San Diego
2019	Doctor of Philosophy in Chemistry, University of California San Diego

PUBLICATIONS

Arismendi-Arrieta, D.J., Riera, M., Bajaj, P., Prosmiti, R., Paesani, F., "i-TTM Model for Ab Initio-Based Ion–Water Interaction Potentials. 1. Halide – Water Potential Energy Functions", *J. Phys. Chem. B*, 120, 1882, **2016**

Riera, M., Götz, A.W., Paesani, F., "i-TTM Model for Ab Initio-Based Ion–Water Interaction Potentials. II. Alkali metal ion – water potential energy functions", *Phys. Chem. Chem. Phys.*, 18, 30334, **2016**

Reddy, S.K., Straight, S.C., Bajaj, P., Pham, C.H., Riera, M., Moberg, D.R., Morales, M.A., Knight, C., Götz, A.W., Paesani, F., "On the accuracy of the MB-pol many-body potential for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice", *J. Chem. Phys.*, 145, 194504, **2016**

Riera, M., Mardirossian, N., Bajaj, P., Götz, A.W., Paesani, F., "Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces", *J. Chem. Phys.*, 147, 161715, **2017**

Riera, M., Brown, S.E., Paesani, F., "Isomeric Equilibria, Nuclear Quantum Effects, and Vibrational Spectra of $M^+(H_2O)_{n=1-3}$ Clusters, with M = Li, Na, K, Rb, and Cs, through Many-Body Representations", *J. Phys. Chem. A*, 122, 5811, **2018**

Zhuang, D., Riera, M., Schenter, G.K., Fulton, J.L, Paesani, F., "Many-Body Effects Determine the Local Hydration Structure of Cs^+ in Solution", *J. Phys. Chem. Lett.*, 10, 406, **2019**

Bajaj. P, Riera, M., Lin, J.K., Mendoza Montijo, Y.E., Gazca, J., Paesani, F., "Halide Ion Microhydration: Structure, Energetics, and Spectroscopy of Small Halide–Water Clusters", *J. Phys. Chem. A*, in press, **2019**

FIELDS OF STUDY

Major Field: Chemistry (Physical Chemistry)

Theoretical and Computational Chemistry
Professor Francesco Paesani

ABSTRACT OF THE DISSERTATION

Understanding Many-Body Effects in Ion Hydration with the MB-nrg Potential Energy Functions

by

Marc Riera Riambau

Doctor of Philosophy in Chemistry

University of California San Diego, 2019

Professor Francesco Paesani, Chair

Water is ubiquitous in nature. The arrangement and behavior of water molecules around solutes, ranging from monoatomic ions such as alkali metal ions, to small organic molecules or proteins, is important to understand the driving forces of many processes and reactions. Computer simulations have become a powerful tool within the last decades, but a unified molecular level picture of the mechanism of hydration is still missing. On one hand, *ab initio* methods are accurate, but the treatable size of the system is relatively small, and their accuracy depends on the level of theory used. On the other hand, force fields allow for the study of large system sizes, but at the cost of accuracy. In this work, we present another approach, the MB-nrg (for many-body

energy) potential energy functions (PEFs), which have the speed of a force field and the accuracy of the current "gold standard" in electronic structure calculations: coupled-cluster with singles, doubles and perturbative triples excitations (CCSD(T)) at the complete basis set limit. In the MB-nrg framework, many-body effects are described through classical polarization, and explicit corrections using permutationally invariant polynomials (PIPs) fitted to reproduce CCSD(T) reference data are added at the two- and three-body level. Focusing on the alkali metal ions, it is shown that adding higher levels of corrections to the classical polarizable model improves the agreement between experiments and our theoretical predictions. Finally, this methodology is extendable to any insulator, i.e. the electrons are not delocalized over the whole system, for which accurate reference energies can be obtained. However, the software infrastructure available right now does not allow the use of this kind of methodology in an efficient way. Consequently, we have also developed a new software infrastructure MBX (for many-body expansion) that enables efficient energy calculations. The combination of MBX with MB-nrg is a powerful tool that allows us to study any complex system, as long as accurate reference energies can be obtained.

Introduction

0.1 The problem of water solutions

Two atoms of hydrogen and one atom of oxygen: H_2O. As simple as this molecule might seem, water has been an object of discussion for years. Its properties when isolated in the gas phase are well known, but water has many unique characteristics in the condensed phase. In comparison with other similar molecules such as methane, ammonia, or sulfur dioxide, water has, for example, an unusually high boiling and melting point, high surface tension, and the solid phase has a lower density than the liquid phase. Most of these anomalies are related to the ability of water to form flexible and extended hydrogen bond networks in the condensed phase.[1,2]

Although water is everywhere in nature, it never comes alone. The largest source of water, the oceans, have many other species dissolved or in suspension, such as monoatomic ions, molecular ions, organic molecules, biomolecules and living microorganisms. Other sources of water, such as lakes or the atmosphere, have similar compounds in other concentrations, and even up to 60% of the human body is water. Thus, electrolyte aqueous solutions play an important role in regulation of living systems,[3,4] growth of atmospheric particles,[5,6] energy resources,[7] and others. Inorganic and organic ions are known to be crucial in biological processes such as maintaining the rigidity and hardness of bones and teeth (Ca^{2+}, PO_4^{3-}, and Na^+),[8] neurological signaling, synapses, transport mechanisms (Na^+, K^+),[9] and many others. Other ions and small molecules such as Br^- and SO_2 are also important in environmental chemical reactions. Halide ions in aerosols contribute to the formation of ozone in the polluted marine boundary layer, while in sea spray aerosols (SSA) particles they are responsible for the degradation of tropospheric

ozone. In the atmosphere, the hydrolysis of acidic gases like SO_2 can produce acid rain.[10–13]

When electrolytes are dissolved in water, they tend to dissociate into ionic species, rearranging the water network and modifying the properties of water around them. Depending on the nature of the molecules and ions, such as their size, charge, and polarizability, the alteration of these properties will vary in different ways.[14–20] As an example, Na^+ and Ca^{2+} have very similar ionic radius, and both ions are present in our body. However, the Na/K pump is selective to sodium and not to calcium, mostly because of the arrangement of the water molecules around both ions.[21] This is one of many examples in which the arrangement of water molecules around ions determines the outcome of a process. Although the mechanism of hydration has been studied for years,[22] a clear picture of the driving forces of the hydration process is still missing, leading to different conclusions depending on the method used, in both experiments and theoretical predictions.[23–33] One aspect that has remained elusive is the ion propensity to air-liquid interfaces. Until the early 2000's, the predominant assumption was that the ions preferred to stay hydrated in the bulk rather than being at the surface. However, molecular dynamics (MD) simulations using more advanced potential energy functions or models that account for the polarization of the atoms predicted larger polarizable ions to have a larger propensity for the interface, while smaller ions remained in the bulk.[34–37] This finding lead to new conclusions and the development of new theories about reaction mechanisms happening in biological membranes and atmospheric aerosols. Ions at the surface can participate in reactive uptakes of other gases that are less soluble in aqueous solutions. Second harmonic generation (SHG) and vibrational sum-frequency generation (VSFG), along with photoelectron spectroscopy, have supported the hypothesis that the larger and more polarizable ions tend to be at the interface.[38–43] However, none of these techniques are truly surface specific, and although they do probe the surface, the depth that they reach is difficult to control. Although there is a general consensus that there is a surface affinity of the most polarizable ions, there is still not a quantitative description of the driving forces responsible for this process,[22,44,45] since experiments are not completely selective and theoretical results depend on the method used. For this, an accurate description of the potential

energy surface (PES) of the system is needed.

0.2 Theoretical toolkit

Computer simulations have become a powerful tool in the last decade, and they have been used to gain insights on the structure and properties of liquids and gas phase systems. Regardless of the simulation method used, a description of the PES of the system is needed. *Ab initio* methods can provide an accurate description of the PES. Ideally, the use of correlated quantum chemistry methods such as configuration interaction (CI)[46] or coupled-cluster (CC)[47,48] methods are preferred, since they recover practically all the exact energy thanks to the explicit treatment of electronic correlation. Currently, coupled-cluster with singles, doubles, and perturbative triples excitations (CCSD(T)) is considered the "gold standard" for computational chemistry calculations. However, this level of theory is only computationally affordable for very small systems, roughly, ten atoms or less. For larger systems, a cheaper method is needed, such as the Møller-Plesset perturbation theory of second order (MP2)[49,50] or density functional theory (DFT).[51–53] These methods enable the treatment of larger systems up to a few hundred atoms, but they are farther from the exact treatment, with errors of up to 10 kcal/mol, which is of the order of magnitude of the most typical non-bonded interactions.[50] It is possible to run MD using *ab initio* methods, known as *ab initio* molecular dynamics (AIMD), but typically the system size is less than a few hundred atoms, and the simulation lengths are shorter than a few hundred picoseconds. Consequently, bulk systems are poorly described, and the sampling might not be long enough to obtain robust thermodynamic properties.

When the system size is of the order of thousands or tens of thousands of atoms, empirical potential energy functions (PEFs) or force fields (FFs) must be used. There are a wide variety of FFs: coarse-grained models,[54–56] atomistic rigid point charge models,[57] flexible point charge models[58] and polarizable models,[23,59–65] ordered from lower to higher computational cost and, generally, accuracy. For the latter, the polarization can be taken into account through different

methods such as Drude oscillators[66,66–68] or calculating the induced dipoles.[37,69,70,70–72] These PEFs are usually fine tuned to reproduce both experimental bulk properties such as densities and thermodynamical properties, and *ab initio* data.

Although huge progress has been made in force field development since the beginning of the millennium, there is still a particularly challenging problem from the theoretical point of view. An accurate description of the PES is not enough to provide a quantitative analysis of all systems; a robust way to introduce nuclear quantum effects in MD simulations and anharmonicity in frequency calculations is also necessary. Modern methods such as path-integral MD (PIMD)[73] or centroid MD (CMD)[74] allow, within their limitations, the inclusion of nuclear quantum effects in MD simulations, while the local monomer (LMon) method enables estimates of anharmonic frequencies within the local mode approximation.[75–78]

0.3 New generation of PEFs

As a consequence of this empirical fitting and simplistic functional forms, accurate description of molecular systems and processes across all phases has been a challenge for decades. In the early 70's, Clementi and coworkers[79] presented a methodology to obtain PEFs entirely from *ab initio* data. However, at the time, computers were severely limited in processing power, and performing thousands of coupled-cluster calculations was not feasable. As computer technology has improved, there has been a continuous effort to derive these types of potential energy surfaces (PESs), which can be transferable across different phases (e.g., CC-pol,[80–82] WHBB,[77,83,84] HBB2-pol,[85,86] and MB-pol[87–89]), enabling computer simulations with high accuracy.[90] All of these PEFs are based on the many-body expansion (MBE) of the energy of a system,[91] which states that the energy of any N-molecule system, which we can call N-mer, can be decomposed as a sum of all the ith-body terms, i.e. one-body (1B), two-body (2B), and so forth, as shown in Eq. 1. This expression is formally exact and is independent of the nature of

the molecule (or group of atoms) that defines each monomer.

$$E_{\mathrm{N}} = \sum_{i=1}^{N} V^{1\mathrm{B}}(i) + \sum_{i<j}^{N} V^{2\mathrm{B}}(i,j) + \sum_{i<j<k}^{N} V^{3\mathrm{B}}(i,j,k) + \cdots + V^{\mathrm{NB}}(1,\ldots,N) \tag{1}$$

Each term can be understood within a classical picture. $V^{1\mathrm{B}}(i)$ corresponds to the monomer deformation energy of monomer i, which is the difference in energy of the monomer in the configuration of the N-mer and the energy of that same monomer at the minimum energy configuration. $V^{2\mathrm{B}}(i,j)$ is the interaction of a group of two monomers i and j, with an analogous definition for $V^{3\mathrm{B}}(i,j,k)$, $V^{4\mathrm{B}}(i,j,k,l)$, and so on. Each of these many-body terms can be recursively calculated with:[92–94]

$$V_I^{\mathrm{nB}} = \sum_{m=1}^{n} (-1)^{n-m} \sum_{J \subset I} E_J^{(m)} \tag{2}$$

were, V_I^{nB} is the *nth*-body energy of the subsystem I, composed of n monomers, of the total system with N monomers, and $E_J^{(m)}$ is the total energy of a given subset J composed by m monomers that also belong to I. Then, the many-body expansion in Eq. 1 can be rewritten as:

$$E_{\mathrm{N}} = \sum_{n=1}^{N} \sum_{I=1}^{\binom{n}{N}} V_I^{nB} \tag{3}$$

which can be related to Eq. 1 through:

$$\sum_{I=1}^{\binom{n}{N}} V_I^{nB} = \sum_{i<j<k<\ldots}^{N} V^{\mathrm{nB}}(i,j,k,\ldots) \tag{4}$$

It has been shown that the MBE converges quickly for most insulators, including water, for which the contribution to the total energy of the many-body terms of higher order than three-body are small, with practically all the contributions coming from the one-, two-, and three-body terms.[95–98] Thus, it is possible to recover a close approximation to the total energy by

computing the low order many-body terms at a high level of theory, and the higher order terms at a lower level. This provides a less computationally expensive alternative to the full system energy calculation while providing a similar result. However, for large systems, although the high order many-body contributions are small, there are many of them, so ignoring them would lead to a relatively large error in the total energy.

There have been several attempts to take advantage of the fast convergence of the MBE for water, already mentioned previously.[77,80–86] However, only MB-pol[87–89] has been proven to reproduce and predict water properties across all the phases: from the gas to the condensed phase.[90,99,100] In MB-pol, the one-body energies are obtained using the PEF in Ref. 101, the two-body term has explicit dispersion energy and permanent electrostatics, and the many-body effects are calculated through classical many-body induction with a modified version of the TTM4-F model.[102,103] On top of this classical polarizable potential, explicit corrections to the two- and three-body terms are added with permutationally invariant polynomials (PIPs)[104] that are fitted to reproduce CCSD(T) reference energies for thousands of dimers and trimers at the complete basis set limit. These corrections are only applied at the short range, since at long range, when the electronic densities are not overlapping, the classical description is exact. The PIP corrections are multiplied by a switch function that smoothly turns off the polynomials at a given cutoff (see Refs. 87, 88, and 89 for more details about the development of MB-pol). This methodology is not exclusive for water, but can also be applied to model the interaction between water and other systems such as halide ions,[105,106] alkali metal ions and industrial and atmospherically relevant gases such as CO_2 and SO_2.

0.4 Summary

This dissertation focuses on the development of TTM-nrg (for Thole type model energy), a classical polarizable PEF compatible with MB-pol, and MB-nrg (for many-body energy) to describe interactions between alkali-metal ions and water, both derived entirely from accurate *ab*

initio data, i.e. CCSD(T) at the complete basis set limit. Chapter 1 will present the development and assessment of TTM-nrg, originally called i-TTM for ion TTM energy.[107] Since the many-body contribution to the energy is calculated in the exact same way as for MB-pol, they are compatible by construction. TTM-nrg mantains the accurate MB-pol PEF for water-water interactions, while it uses a simple classical form (Born-Mayer potential) to describe the short range contributions. These PEFs are at least as accurate as DFT, but with several orders of magnitude lower in computatinal cost. In this chapter, the energies and vibrational frequencies, calculated with TTM-nrg, of multiple cluster sizes are compared for each alkali-metal ion to the same properties computed with DFT and other force fields. It is shown that while the errors from DFT highly depend on the ion and the functional used, TTM-nrg has a constant average error per molecule.

In Chapter 2, the two-body term of the MB-nrg PEF for alkali-metal ion water interactions is presented, explaining in detail from the training set generation to the fitting procedure. The chapter also presents comparisons of energies and normal mode frequencies between $M^+(H_2O)$ dimers, with M = Li, Na, K, Rb, and Cs, across different DFT functionals, classical polarizable models, coupled-cluster and MB-nrg, showing that MB-nrg is the closest model to the "gold standard" CCSD(T) at the complete basis set limit. This chapter shows the improvement due to the replacement of the classical Born-Mayer repulsion expression with the permutationally invariant polynomials, being able to predict energies of a dimer within 0.1 kcal/mol and deviations in the frequencies of a maximum of 30 wavenumbers.

After showing that MB-nrg reproduces coupled-cluster energies and frequencies, Chapter 3 presents an example of application of MB-nrg to predict infrared (IR) spectra of small $M^+(H_2O)_n$ gas phase clusters, with n=1-3. Using replica-exchange molecular dynamics (REMD), the different isomers present in the gas phase for each metal ion along their relative fraction at each temperature are obtained. The calculated anharmonic temperature dependent IR spectra are then compared with experimental spectra, showing a good agreement between the two data sets.

Since the interaction between Cs^+ and water is the weakest of all the alkali metal ions,

there is a competition between water-water and water-ion interactions. Thus, this ion is taken as a case study to unravel the effect of many-body interactions on the hydration structure of electrolytes. Chapter 4 presents a comparison of the hydration structure using four different models: a point charge, non-polarizable model (TIP4P),[108,109] a classical polarizable model (TTM-nrg), and two models able to correct for the quantum effects: MB-nrg with corrections to the two-body term (MB-nrg 2B+NB), and MB-nrg with corrections in both the two-body and the three-body terms (MB-nrg 2B+3B+NB). This chapter shows that, although there are small deviations in the many-body contributions to the energy among the different methods used, these small effects can completely change the hydration structure of the ion. Thus, having a correct description of these interactions is key to understand the driving forces of ion hydration.

Chapter 5 presents the software developed to automatize and generalize the process of obtaining TTM-nrg and MB-nrg PEFs and facilitate their usage. It first shows a general overview of the fitting protocol, from the generation of configurations, through the fitting process, to the final implementation in the energy software. The second part of the chapter introduces the basic ideas and structure of MBX, the energy calculator. In the near future both parts of the software will be combined and publically released.

The final chapter of this thesis, Chapter 6, summarizes the main conclusions of this work. We have shown that it is possible to obtain PEFs with an unprecedented accuracy that are as computationally efficient as a polarizable force field, and that they can be systematically improved by adding higher order correction terms to the different many-body contributions to the energy.

Chapter 1

i-TTM Model for Ab Initio-Based Ion-Water Interaction Potentials. Alkali Metal Ion–Water Potential Energy Functions.

1.1 Introduction

Electrolyte solutions play important roles in numerous phenomena ranging from protein stability inside living cells to formation of atmospheric aerosol particles and electrochemical processes at electrode interfaces.[14–16] Aqueous solutions that contain alkali metal ions are particularly important in biochemistry due to the regulative role they have in cell physiology. Na^+/K^+ pumps, for instance, generate electrochemical potentials across cell membranes, which play crucial roles in nerve signal transmission.[17] Ions at aqueous interfaces of aerosol particles also appear to be directly involved in several physicochemical processes related to new particle formation and growth in the atmosphere.[18–20] Furthermore, aqueous rechargeable Li^+ and Na^+ batteries are particularly promising for large-scale energy storage since they are not flammable, use inexpensive salts, can conduct ions more effectively than organic electrolytes, and are based on environmentally benign electrolytes.[110]

[0†] Electronic Supplementary Information (ESI) available: Analysis of the basis set convergence and description of the extrapolation procedure for the CCSD(T) and CCSD(T)-F12b calculations. See DOI: 10.1039/b000000x/

[0a] *Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, U.S.A. Email: fpaesani@ucsd.edu; Telephone: +1 858 - 822 - 3383*

[0b] *San Diego Supercomputer Center, University of California, San Diego, La Jolla, California 92093, U.S.A.*

[0‡] To whom correspondence should be addressed.

Both experimental and theoretical studies aimed at characterizing the microscopic mechanisms of ion hydration have been reported in the literature.[22] However, despite significant progress, it remains unclear how the local hydration structure leads to different solution properties.[23–33] With the development of efficient *ab initio* molecular dynamics approaches[22,59,111] and polarizable force fields,[23,59–65] computer simulations have become a powerful tool to investigate the properties of electrolyte solutions at the molecular level.

The starting point to achieve a molecular-level understanding of the hydration of alkali metal ions (M^+) in different environments is the investigation of the properties of small $M^+(H_2O)_n$ clusters. Several polarizable force fields have been developed in the last decade to model the behavior of both halide and alkali metal ions in aqueous environments.[60,63,64,68,112–119] However, most of these force fields treat the water molecules as rigid, which precludes comparisons with vibrational spectra that directly probe the structural and dynamical properties of different hydration shells.[120–123] A remarkable exception is the AMOEBA force field, which includes both intramolecular flexibility and many-body effects through classical polarization.[63,124] In the past years, there has been extensive work on developing AMOEBA parameterizations for metal ion-water interactions, including alkali metals,[124] alkaline-earth metals,[125] transistion metals,[126,127] and lanthanides and actinides.[128] Since these AMOEBA parameterizations were performed by targeting condensed phase systems, they primarily aim at reproducing bulk properties (e.g., radial distribution functions and hydration free energies) of electrolyte solutions.

Beginning with the pioneering work by Clementi and coworkers,[79] there has been a long standing effort to derive potential energy surfaces (PESs) entirely from *ab initio* data, which can be transferable across different phases. The present study describes the extension of the i-TTM (ion-Thole-type model) scheme, introduced in Ref. 105 for describing halide ion-water interactions, to the representation of *ab initio*-based alkali metal ion-water PESs compatible with the many-body MB-pol potential for water.[87–89,129] The new i-TTM PESs are derived from fits to coupled-cluster electronic structure data and build upon an extended Thole-type polarizable model[102] that takes explicitly into account the mutual polarization between the alkali metal ions

and the water molecules. The article is organized as follows: The functional form of the i-TTM PESs is described in Section 2. The accuracy of the i-TTM PESs is then assessed in Section 3 through comparisons with both *ab initio* and force field results obtained for small $M^+(H_2O)_n$ clusters, with $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+, and $n = 1-4$. A summary and outlook on future work are given in Section 4.

1.2 Model development and computational details

1.2.1 Electronic structure calculations

All reference energies for the $M^+(H_2O)$ dimers were calculated using coupled cluster theory with single, double and perturbative triple excitations, CCSD(T), or explicitly correlated CCSD(T)-F12b theory,[130,131] depending on the alkali metal ion. The calculations were performed using triple-ζ and quadruple-ζ basis sets and the resulting energies extrapolated to the complete basis set (CBS) limit via a two-point approximation (see Supplementary Information).[132,133] For $Li^+(H_2O)$ and $Na^+(H_2O)$, CCSD(T)-F12b was used with the aug-cc-pVnZ (n = T, Q) basis sets for oxygen and hydrogen and the cc-pCVnZ (n = T, Q) basis sets for lithium and sodium.[134–136] Since no suitable basis sets for CCSD(T)-F12b calculations are available for the heavier alkali metals, CCSD(T) calculations were performed for ion-water clusters containing potassium, rubidium, and cesium. For $K^+(H_2O)$ the aug-cc-pVnZ (n = T, Q) basis sets were used for oxygen and hydrogen, and the cc-pwCVnZ (n = T, Q) basis set for potassium.[137] Correlation consistent basis sets are not available for rubidium and cesium, hence the def2-nZVPP (n = T, Q) basis sets[138] were used for $Rb^+(H_2O)$ and $Cs^+(H_2O)$. These basis sets include effective core potentials with 28 and 46 electrons for Rb^+ and Cs^+, respectively. The Boys-Bernardi counterpoise correction[139] for the basis set superposition error (BSSE) was applied to all calculations carried out with the def2-nZVPP basis sets. The optimal combination of methods and basis sets for the different alkali metals was determined through careful analysis of the convergence of the results obtained with both CCSD(T) and CCSD(T)-F12b (see Supplementary

Table 1.1. Electronic structure methods and basis sets employed for the calculation of $M^+(H_2O)$ dimer reference energies

Dimer	Basis set (n = T, Q)		Method
	O, H	M	
$Li^+(H_2O)$	aug-cc-pVnZ	cc-pCVnZ	CCSD(T)-F12b
$Na^+(H_2O)$	aug-cc-pVnZ	cc-pCVnZ	CCSD(T)-F12b
$K^+(H_2O)$	aug-cc-pVnZ	cc-pwCVnZ	CCSD(T)
$Rb^+(H_2O)$	def2-nZVPP	def2-nZVPP	CCSD(T)
$Cs^+(H_2O)$	def2-nZVPP	def2-nZVPP	CCSD(T)

Information). Table 1.1 summarizes the combination of electronic structure methods and basis sets used for the reference calculations in this study. In addition, density fitted MP2 calculations (DF-MP2) were performed with the same basis sets to assess the accuracy of the i-TTM potentials. All DF-MP2, CCSD(T), and CCSD(T)-F12b calculations were carried out with MOLPRO.[140,141]

Following Ref. 105, the *ab initio* energies were computed for $M^+(H_2O)$ dimer configurations along radial scans corresponding to different orientations of M^+ relative to H_2O. In these scans, the distance between M^+ and the water oxygen was set to vary from 1.0 to 12.0 Å for $Li^+(H_2O)$ and $Na^+(H_2O)$, and from 1.5 to 12.0 Å for all other alkali metal ion-water dimers, with the polar angle θ varying from 0° to 90° in intervals of 45° and the azimuthal angle ϕ being set to 0°, 52.345°, 90°, and 180°. In these dimer calculations, the water molecule was kept fixed at the vibrationally averaged geometry, with r_{OH} = 0.9716257 Å and θ_{HOH} =104.69°.[142] In total, ∼450 dimer configurations were included in the training sets used to fit each i-TTM potential.

As in Ref. 143, the dipole polarizability of each alkali metal ion, α_{M^+}, was computed at the CCSD(T)(-F12b) level of theory using the basis sets aug-cc-pCV5Z for Li^+ and Na^+, aug-cc-pwCV5Z for K^+, and def2-QZVPP for Rb^+ and Cs^+. The values of α_{M^+} obtained for Li^+, Na^+, K^+, Rb^+, and Cs^+ are 0.0285 $Å^3$, 0.1476 $Å^3$, 0.8184 $Å^3$, 1.2758 $Å^3$, and 2.3377 $Å^3$, respectively, in agreement with other theoretical estimates previously reported in the literature.[144,145]

The dispersion coefficients C_{6,OM^+} and C_{6,HM^+} were calculated using the exchange-hole dipole model (XDM) developed by Becke and Johnson.[146] The XDM calculations were carried

out with the Postg software[147,148] employing electron densities obtained for large alkali metal ion-water separations with the LC-ωPBE[149–151] functional in combination with the aug-cc-pV5Z basis set for $M^+ = Li^+$, Na^+ and K^+, and the def2-QZVPP basis set for $M^+ = Rb^+$ and Cs^+. All DFT calculations presented in the following sections were carried out with Gaussian 09.[152]

1.2.2 AMOEBA calculations

All AMOEBA calculations were performed using TINKER,[153,154] version 6.3, with the AMOEBA 2009 force field,[155] which includes the 2003 parameterizations for both water and alkali metal ions.[124] Although a new version of the AMOEBA force field for water has recently been developed,[156] updated parameterizations for alkali metal ion-water interactions compatible with the new water model are not yet available.[157]

1.2.3 i-TTM model

The total energy, V_{tot}, of a system containing a single alkali metal ion M^+ and N water molecules can be formally decomposed into a sum of three distinct terms describing the energy contributions associated with the intramolecular distortions of each water molecule, V_w^{intra}, the intermolecular water-water interactions, V_w^{inter}, and the ion-water intermolecular interactions, V^{i-TTM},

$$V_{tot} = V_w^{intra} + V_w^{inter} + V^{i-TTM} \tag{1.1}$$

In Eq. 1.1, V_w^{intra} corresponds to the Partridge-Schwenke PES[101] and V_w^{inter} is represented by the MB-pol many-body water potential.[87–89,129] The interested reader is referred to a recent review for detailed comparisons between MB-pol and modern polarizable and DFT models, which show that MB-pol currently is the most accurate potential energy function available for modeling the properties of water from the dimer in the gas phase to the liquid phase and ice.[90] V^{i-TTM}, which was introduced in Ref. 105 for halide ion-water systems, is given by a sum of

four terms

$$V^{\mathrm{i-TTM}} = V^{\mathrm{TTM,elec}} + V^{\mathrm{TTM,ind}} + \sum_{i=1}^{n} \left(V_i^{\mathrm{rep}} + V_i^{\mathrm{disp}} \right) \tag{1.2}$$

Here, $\mathrm{V}^{\mathrm{TTM,elec}}$ and $\mathrm{V}^{\mathrm{TTM,ind}}$ represent the electrostatic interactions between permanent point charges and induced dipole moments on the water molecules and the ion, respectively.[87,105] An extended Thole-type model, originally introduced with the TTM4-F water potential,[102] is adopted to represent induction interactions using dipole polarizabilities determined from electronic structure calculations as described above. The repulsive potential, V_i^{rep}, is represented by a sum of pairwise Born-Mayer functions between the O and H atoms of water molecule i and the alkali metal ion M^+,

$$\begin{aligned} V_i^{\mathrm{rep}} =& A_{\mathrm{OM}^+} e^{b_{\mathrm{OM}^+} R_{\mathrm{O}_i\mathrm{M}^+}} \\ &+ A_{\mathrm{HM}^+} e^{b_{\mathrm{HM}^+} R_{\mathrm{H}_{1,i}\mathrm{M}^+}} + A_{\mathrm{HM}^+} e^{b_{\mathrm{HM}^+} R_{\mathrm{H}_{2,i}\mathrm{M}^+}} \end{aligned} \tag{1.3}$$

In Eq. 1.3, $R_{\alpha\beta}$ are the interatomic distances between α = O, H and β = M^+, while $A_{\alpha\beta}$ and $b_{\alpha\beta}$ are fitting parameters.

The dispersion energy V_i^{disp} in Eq. 1.2 is also represented through a sum of pairwise terms

$$\begin{aligned} V_i^{\mathrm{disp}} =& - f_6(R_{\mathrm{O}_i\mathrm{M}^+}, \delta_{\mathrm{OM}^+}) \frac{C_{6,\mathrm{OM}^+}}{R_{\mathrm{O}_i,\mathrm{M}^+}} \\ &- f_6(R_{\mathrm{H}_{1,i}\mathrm{M}^+}, \delta_{\mathrm{HM}^+}) \frac{C_{6,\mathrm{HM}^+}}{R_{\mathrm{H}_{1,i},\mathrm{M}^+}} \\ &- f_6(R_{\mathrm{H}_{2,i}\mathrm{M}^+}, \delta_{\mathrm{HM}^+}) \frac{C_{6,\mathrm{HM}^+}}{R_{\mathrm{H}_{2,i},\mathrm{M}^+}} \end{aligned} \tag{1.4}$$

where $f_6(R, \delta)$ are Tang–Toennies damping functions[158] with damping coefficients δ, and C_{6,OM^+} and C_{6,HM^+} are the dispersion coefficients obtained using the XDM method as described in Section 1.2.1.

1.2.4 Fitting procedure

The i-TTM parameters A_{OM^+}, A_{HM^+}, b_{OM^+}, and b_{HM^+} of V^{rep} as well as δ_{OM^+} and δ_{HM^+} of V^{disp} were optimized by a fitting procedure in which the linear parameters were determined through singular value decomposition while the nonlinear parameters were obtained using the simplex algorithm. Following Ref. 158, the parameters δ_{OM^+} and δ_{HM^+} of the Tang-Toennies damping functions were constrained to be equal to the parameters b_{OM^+} and b_{HM^+} of the corresponding Born-Mayer functions since the damping of both $\mathrm{V}^{\mathrm{disp}}$ and V^{rep} has the same physical origin, i.e., the overlap of the electron densities of the two monomers (M^+ and $\mathrm{H_2O}$).

In the fitting procedure, the weighted sum of the squared residuals, $\chi^2 = \sum_m w_m[(V_{\mathrm{i-TTM}}(m) - E^{\mathrm{ref}}(m)]^2$, was minimized with respect to the CCSD(T)(-F12b) reference interaction energies E^{ref}. The weights w_m were chosen to emphasize dimers with low total energies according to $w(E) = (\frac{E_{\mathrm{weight}}}{E - E_{\mathrm{min}} + E_{\mathrm{weight}}})^2$, where E_{min} is the $\mathrm{M}^+(\mathrm{H_2O})$ global minimum interaction energy and E_{weight} corresponds to the range of favorably weighted energies. Specifically, $E_{\mathrm{weight}} = 50$ kcal/mol for $\mathrm{Li}^+(\mathrm{H_2O})$ and $E_{\mathrm{weight}} = 40$ kcal/mol for $\mathrm{Na}^+(\mathrm{H_2O})$, $\mathrm{K}^+(\mathrm{H_2O})$, $\mathrm{Rb}^+(\mathrm{H_2O})$ and $\mathrm{Cs}^+(\mathrm{H_2O})$.

The root-mean-square deviations (RMSDs) associated with the fits are 1.48 kcal/mol for $\mathrm{Li}^+(\mathrm{H_2O})$, 1.06 kcal/mol for $\mathrm{Na}^+(\mathrm{H_2O})$, 0.60 kcal/mol for $\mathrm{K}^+(\mathrm{H_2O})$, 0.60 kcal/mol for $\mathrm{Rb}^+(\mathrm{H_2O})$, and 0.77 kcal/mol for $\mathrm{Cs}^+(\mathrm{H_2O})$. All alkali metal ion i-TTM parameters are listed in Table 1.2.

1.3 Results

1.3.1 $\mathrm{M}^+(\mathrm{H_2O})$ dimers

To assess the accuracy of the new alkali metal ion-water potentials, in Table 1.4 the i-TTM interaction energies of the five $\mathrm{M}^+(\mathrm{H_2O})$ dimers at their minimum geometries (obtained at the DF-MP2/CBS level of theory) are compared with the corresponding CCSD(T)(-F12b)/CBS and DF-MP2/CBS values. The i-TTM potentials agree well with the most accurate CCSD(T)(-F12b)

Table 1.2. Parameters for the i-TTM alkali metal ion-water potentials defined in Equation 1.2.

i-TTM term[a]	Parameter	$Li^+(H_2O)$	$Na^+(H_2O)$	$K^+(H_2O)$	$Rb^+(H_2O)$	$Cs^+(H_2O)$
V^{rep}	A_{OM^+} (kcal/mol)	32318.00	47827.70	49780.20	68268.10	45030.70
	A_{HM^+} (kcal/mol)	3245.78	4992.61	4941.43	6982.09	9426.17
	b_{OM^+} ($Å^{-1}$)	4.02327	3.76951	3.39759	3.35667	3.03294
	b_{HM^+} ($Å^{-1}$)	4.00672	3.82261	3.31969	3.29901	3.26732
$V^{TTM,ind}$	α_{M^+} ($Å^3$)	0.0285	0.1476	0.8184	1.2758	2.3377
V^{disp}	C_{6,OM^+} ($Å^6$ kcal/mol)	43.147	176.255	431.432	527.192	700.097
	C_{6,HM^+} ($Å^6$ kcal/mol)	18.712	85.787	218.843	271.321	365.322

[a] In the $V^{TTM,ind}$ term, the polarizabilities of the O and H atoms of the water molecules and associated damping parameters correspond to the MB-pol values defined in Ref. 87.

reference energies, with deviations that are smaller than 1 kcal/mol in all cases. Importantly, the accuracy of the i-TTM interaction energies is comparable to that of DF-MP2, which exhibits similar deviations from the CCSD(T)-(F12b) reference values. Figure 1.1 shows a comparison between the i-TTM and CCSD(T)(-F12b)/CBS potential energy curves for alkali metal ion-water dimers as a function of the distance R_{OM^+} between M^+ and the water oxygen for different (θ, ϕ) orientations of M^+ relative to H_2O. As expected from the RMSD values associated with

Table 1.3. Comparison between the interaction energies (in kcal/mol) for $M^+(H_2O)$ dimers calculated at the CCSD(T)-(F12b)/CBS, DF-MP2/CBS, and i-TTM levels of theory. The dimer structure has been optimized at DF-MP2 level of theory using a triple ζ basis set

Ion	CCSD(T)(-F12b)/CBS	DF-MP2/QZ	i-TTM
$Li^+(H_2O)$	-34.81	-34.43	-35.28
$Na^+(H_2O)$	-24.14	-23.89	-24.86
$K^+(H_2O)$	-17.97	-18.13	-17.02
$Rb^+(H_2O)$	-15.47	-15.77	-14.82
$Cs^+(H_2O)$	-13.90	-14.35	-12.54

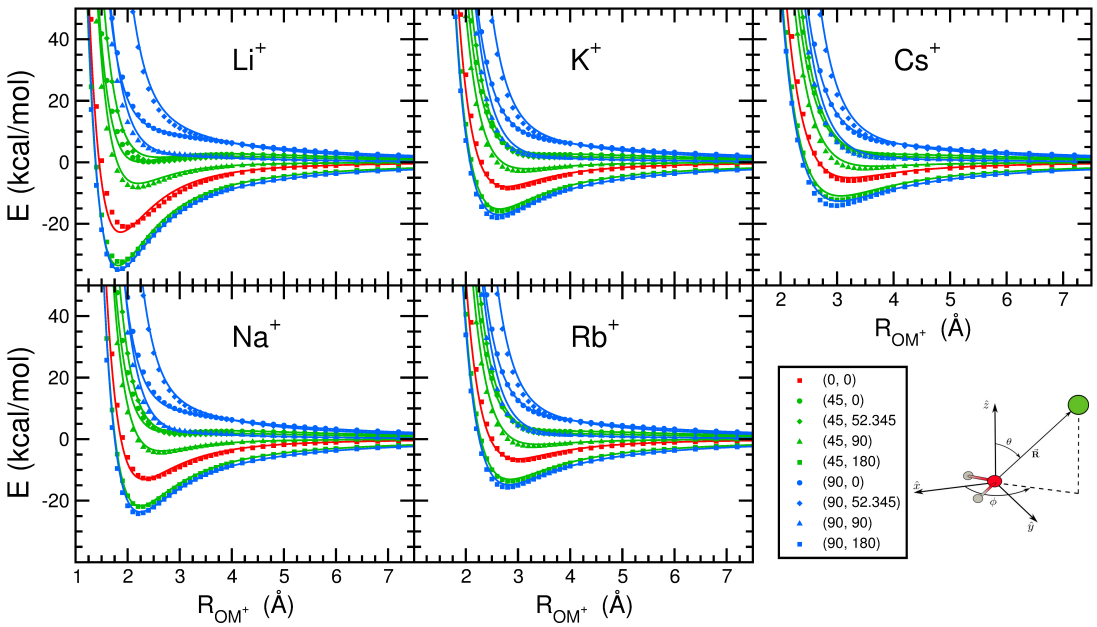

Figure 1.1. i-TTM potential energy scans (solid lines) and corresponding CCSD(T)(-F12b)/CBS reference data (symbols) for $M^+(H_2O)$ dimers with $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+. The numbers in parenthesis indicate different (θ, ϕ) orientations (in degrees) of M^+ relative to H_2O within the coordinate frame defined in Ref. 105 and shown in the lower right panel.

the fits, the i-TTM potentials correctly reproduce both anisotropy and relative strength of the different alkali metal ion-water interactions. The largest RMSD obtained for the $Li^+(H_2O)$ dimer is reflected in somewhat larger deviations between i-TTM and CCSD(T)-F12b/CBS values in the short range of the potential energy scans as compared to the other alkali metal ions. The larger error in the $Li^+(H_2O)$ dimer i-TTM PES is a consequence of the correlation between the strength of the ion-water interaction and the radius of the alkali metal cations. Since Li^+ is much smaller than the other alkali metal cations, it is the strongest Lewis acid, which results in stronger (and more covalent-like) interactions between Li^+ and H_2O. Due to the overlap of Li^+ and water electron densities at short distances, these interactions are more quantum-mechanical in nature and cannot be completely recovered by classical potential energy functions such as $V^{\mathrm{i-TTM}}$. Similar deviations from the reference values were also found in the i-TTM description of the interactions between F^- and H_2O.[105] As demonstrated in Ref. 106 for the halide ion-water interactions, these limitations associated with a classical representation of the

interactions, can be overcome through a rigorous description of ion-water two-body effects at short range. Figure 1.2 shows a comparison between the potential energy angular profiles (PEAPs) calculated with the i-TTM potentials for the five alkali metal ions and the corresponding CCSD(T)(-F12b)/CBS reference values. Also shown are the results obtained with the AMOEBA 2009 force field. The PEAPs are computed along the azimuthal angle ϕ for planar configurations ($\theta = 90^\circ$) of each dimer keeping the water molecule fixed at its vibrationally averaged geometry. In these calculations the distance R_{OM^+} was fixed to the minimum energy distance of each dimer for $\phi = 180^\circ$ and $\theta = 90^\circ$, i.e., $R_{OLi^+} = 1.824$ Å, $R_{ONa^+} = 2.209$ Å, $R_{OK^+} = 2.594$ Å, $R_{ORb^+} = 2.787$ Å, and $R_{OCs^+} = 2.969$ Å. All PEAPs display two symmetric energy maxima at $\phi \approx 55^\circ$ and $\phi \approx 305^\circ$, corresponding to geometries at which the water hydrogen atoms point towards the alkali metal ion.

The i-TTM PEAPs closely follow the corresponding CCSD(T)(-F12b)/CBS curves for all alkali metal ion-water dimers. Somewhat larger deviations from the reference values are found in the i-TTM PEAP of the $Li^+(H_2O)$ dimer, particularly in the repulsive regions. As mentioned above, improving upon the i-TTM results would require an explicit account for quantum-mechanical effects at short range, such as charge transfer/penetration and Pauli repulsion, which cannot be fully represented through classical expressions. Although AMOEBA describes all PEAPs reasonably well in regions near the global minimum ($\phi = 180^\circ$), large deviations relative to the reference data are found in the repulsive regions. In particular, AMOEBA underestimates the maxima in the $Li^+(H_2O)$ and $Na^+(H_2O)$ PEAPs by approximately 60 kcal/mol and 20 kcal/mol, respectively.

1.3.2 Small $M^+(H_2O)_n$ clusters

To assess the transferability of the i-TTM potentials for alkali metal ions to systems with several water molecules, the i-TTM interaction energies calculated for small $M^+(H_2O)_n$ clusters with $n = 1-4$ are compared in Table 1.5 with the corresponding CCSD(T)(-F12b), DF-MP2, and AMOEBA values. As in Table 1.4, the interaction energies are defined as the total energy of

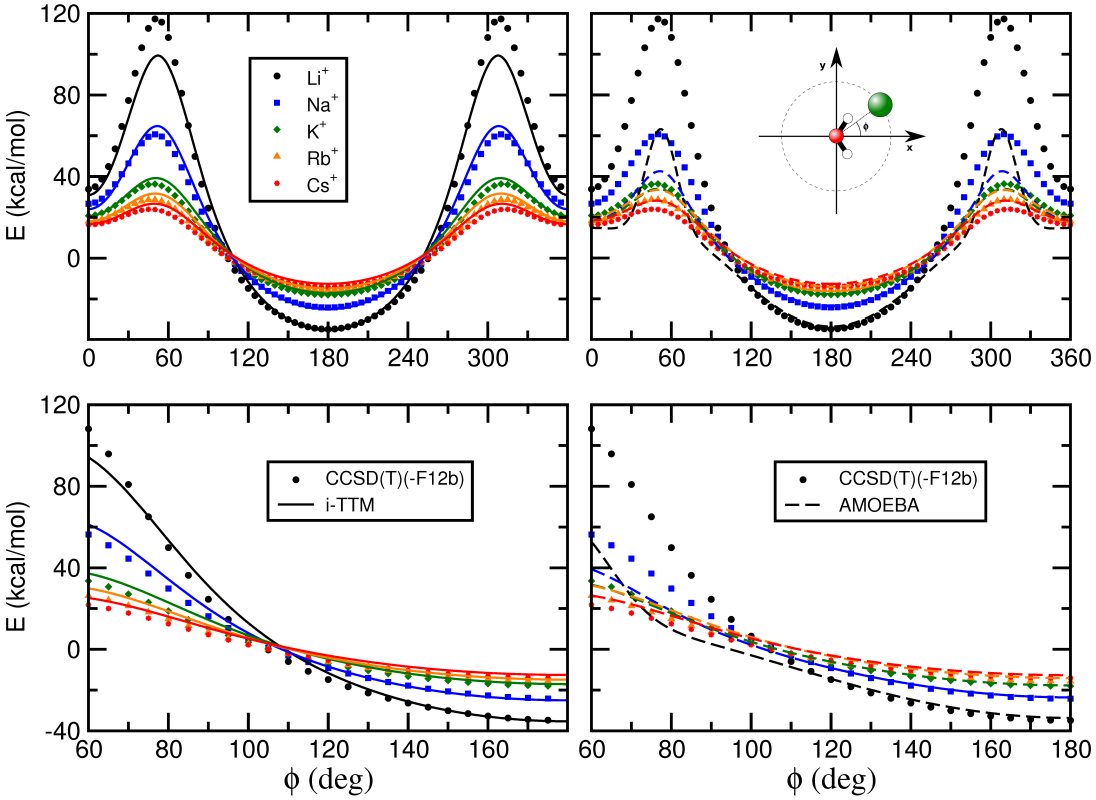

Figure 1.2. Potential energy angular profiles (PEAPs) on the $M^+(H_2O)$ dimer PESs with M^+ = Li^+, Na^+, K^+, Rb^+, and Cs^+ calculated as a function of ϕ for $\theta = 90°$ (see Figure 1.1) for R_{OM^+} equal to the minimum energy distance of each dimer, keeping the water molecule fixed at its vibrationally averaged geometry. The i-TTM results are shown as solid lines (left), the AMOEBA results as dashes lines (right), and the CCSD(T)(-F12)/CBS reference values as filled symbols.

Table 1.4. Comparison between the interaction energies (in kcal/mol) for $M^+(H_2O)$ dimers calculated at the CCSD(T)-(F12b)/CBS, DF-MP2/CBS, and i-TTM levels of theory. The dimer structure has been optimized at DF-MP2 level of theory using a triple ζ basis set

Ion	CCSD(T)(-F12b)/CBS	DF-MP2/QZ	i-TTM
$Li^+(H_2O)$	-34.81	-34.43	-35.28
$Na^+(H_2O)$	-24.14	-23.89	-24.86
$K^+(H_2O)$	-17.97	-18.13	-17.02
$Rb^+(H_2O)$	-15.47	-15.77	-14.82
$Cs^+(H_2O)$	-13.90	-14.35	-12.54

the cluster (optimized at the DF-MP2/aug-cc-pVTZ level of theory) minus the energies of the individual water molecules (in the distorted configurations within the cluster). This comparison thus directly probes the actual intermolecular interactions without being affected by differences in the representation of the water distortion energies due to different PESs. To approach the complete basis set limit, the CCSD(T)(-F12b) values were obtained through a many-body decomposition of the interaction energy according to the SAMBA algorithm described in Ref. 132. For Li^+, Na^+, and K^+ clusters, the two-body interaction energies were calculated as described in Ref. 132 through a two-point extrapolation of the corresponding values obtained with the aug-cc-pVTZ and aug-cc-pVQZ basis sets. The three-body contributions were obtained using the aug-cc-pVTZ basis set with BSSE corrections, and four- and five-body contributions were calculated at the CCSD(T)(-F12b)/aug-cc-pVTZ level of theory. For Rb^+ and Cs^+ clusters, due to the smaller size of the def2-nZVPP basis sets, both two-body and three-body contributions were obtained through a two-point extrapolation of the def2-TZVPP and def2-QZVPP results, including BSSE corrections. The four- and five-body contributions were calculated using CCSD(T)/def2-QZVPP with BSSE corrections. It was shown in Ref. 132 that this approach currently provides the most accurate interaction energies for molecular clusters at the CCSD(T) level. Due to the lower computational cost, the DF-MP2 interaction energies were instead calculated for the full clusters using the aug-cc-pVQZ (def2-QZVPP for K^+, Rb^+, and Cs^+) basis set at all orders.

Table 1.5. CCSD(T)(-F12b), DF-MP2, i-TTM, and AMOEBA interaction energies (in kcal/mol) for $M^+(H_2O)_n$ clusters, with $n = 1-4$ and M^+ = Li^+, Na^+, K^+, Rb^+, and Cs^+. Also reported are the deviations (δE, in kcal/mol) of DF-MP2, i-TTM and AMOEBA from the CCSD(T)-(F12b) values[a].

Alkali metal ion	n	(Symmetry)	CCSD(T)(-F12b)	DF-MP2	δE	i-TTM	δE	AMOEBA	δE
Li^+	1	(C_{2v})	-34.81	-34.43	-0.38	-35.28	0.47	-33.41	-1.40
	2	(D_{2d})	-64.74	-64.12	-0.61	-62.94	-1.79	-61.52	-3.22
	3	(D_3)	-88.51	-87.59	-0.92	-81.20	-7.32	-82.13	-6.38
	4	(S_4)	-106.05	-104.94	-1.10	-93.80	-12.25	-97.47	-8.58
AAE[b] per water molecule			–	–	0.32	–	1.72	–	1.82
Na^+	1	(C_{2v})	-24.14	-23.89	-0.25	-24.86	0.73	-23.44	-0.70
	2	(D_{2d})	-45.74	-45.33	-0.41	-46.22	0.48	-44.47	-1.27
	3	(D_3)	-63.96	-63.38	-0.58	-63.12	-0.84	-61.84	-2.12
	4	(C_2)	-79.28	-78.89	-0.39	-78.64	-0.64	-76.44	-2.84
AAE[b] per water molecule			–	–	0.19	–	0.35	–	0.69
K^+	1	(C_{2v})	-17.97	-18.13	0.16	-17.02	-0.95	-17.63	-0.34
	2	(D_{2d})	-33.89	-34.18	0.29	-32.05	-1.84	-33.27	-0.62
	3	(C_2)	-49.25	-49.84	0.59	-46.89	-2.36	-47.75	-1.60
	4	(C_2)	-62.66	-63.43	0.77	-59.33	-3.32	-61.04	-1.62
AAE[b] per water molecule			–	–	0.17	–	0.87	–	0.40
Rb^+	1	(C_{2v})	-15.47	-15.77	0.30	-14.82	-0.65	-14.11	-1.36
	2	(C_{2d})	-29.23	-29.83	0.60	-28.03	-1.20	-26.89	-2.34
	3	(C_2)	-44.10	-45.01	0.91	-42.48	-1.62	-40.69	-3.41
	4	(C_2)	-55.84	-57.03	1.19	-53.63	-2.21	-51.63	-4.20
AAE[b] per water molecule			–	–	0.30	–	0.59	–	1.18
Cs^+	1	(C_{2v})	-13.90	-14.35	0.45	-12.54	-1.36	-12.66	-1.23
	2	(C_s)	-26.66	-27.57	0.92	-24.03	-2.62	-23.55	-3.11
	3	(C_2)	-40.90	-42.12	1.21	-37.81	-3.09	-37.76	-3.14
	4	(C_2)	-51.16	-52.88	1.72	-47.25	-3.91	-47.54	-3.62
AAE[b] per water molecule			–	–	0.44	–	1.17	–	1.18

[a] All calculations were carried out for cluster geometries optimized at the DF-MP2/aug-cc-pVQZ (M^+ = Li^+, Na^+) and DF-MP2/def2-QZVPP (M^+= K^+, Rb^+, Cs^+) level of theory. See the main text for details. [b] Average Absolute Error per water molecule.

The results reported in Table 1.5 indicate that DF-MP2 reproduces the interaction energies for $M^+(H_2O)_n$ clusters with $M^+ = Na^+$ and K^+ to within 1 kcal/mol of the CCSD(T)(-F12b) reference values, with an average absolute error (AAE) per water molecule under 0.2 kcal/mol. The deviations from the CCSD(T) reference data are larger for Li^+ and the heavier alkali metal ions, exceeding an AAE per water molecule of 0.4 kcal/mol in the case of Cs^+. The i-TTM potentials show a slightly worse agreement with the reference data than DF-MP2, with notably larger deviations for the Cs^+ clusters and the larger Li^+ clusters. The accuracy of the AMOEBA force field is comparable to i-TTM. Both i-TTM and AMOEBA show significant deviations for the Li^+ clusters, with the error per water molecule increasing significantly with increasing cluster size, leading to an AAE per water molecule of 1.7 kcal/mol and 1.8 kcal/mol, respectively. This indicates that there are substantial many-body effects that are not captured by these polarizable force fields, which reinforces the notion that a purely classical representations of induction interactions is likely not sufficient to achieve chemical accuracy in the description of alkali metal ion-water interactions.

The minimum energy structures for the $M^+(H_2O)_n$ clusters with $n = 1-4$ are shown in Figure 1.3 and labeled according to the notation (m+m′) where m is the number of water molecules directly bonded to the cation and m′ is the number of water molecules found in the outer hydration shells. Table 1.6 lists the symmetry of the minimum energy structure for all clusters. For all clusters, the i-TTM potentials predict symmetries in agreement with the DF-MP2 results and previous MP2 studies.[159] As shown in Figure 1.3, different coordination topologies and hydrogen bonding patterns start emerging as the cluster size increases, with the H_2O molecules either forming a single hydration shell around M^+ (e.g., see the (3+0) and (4+0) configurations for $Li^+(H_2O)_n$ with $n = 3,4$) or developing a multi-shell structure (e.g., (2+1) and (3+1) configurations for $Cs^+(H_2O)_n$ with $n = 3,4$). For larger ions, hydrogen bonding becomes more important, favoring the location of the ions at the surface as the cluster size increases.

To provide quantitative insights into the interplay and competition between ion-water and water-water interactions in determining the energetics of alkali metal ion-water systems,

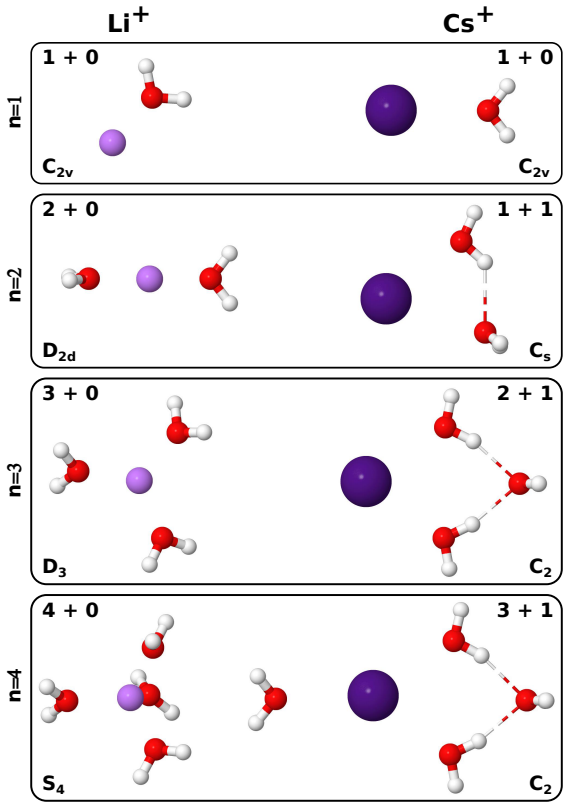

Figure 1.3. Optimized structures for $M^+(H_2O)_n$ clusters with $M^+ = Li^+$ (left, small ion), and Cs^+ (right, large ion) for $n = 1-4$ calculated with DF-MP2/aug-cc-pVTZ and DF-MP2/def2-TZVPP, respectively. The figures were prepared using Jmol.[160]

Table 1.6. Symmetry of the minimum energy clusters for $M^+(H_2O)_n$, with $M^+ = Li^+$, Na^+, K^+, Rb^+ and Cs^+, and $n = 1-4$.

Cluster Size	$Li^+(H_2O)_n$	$Na^+(H_2O)_n$	$K^+(H_2O)_n$	$Rb^+(H_2O)_n$	$Cs^+(H_2O)_n$
$n=1$	C_{2v}	C_{2v}	C_{2v}	C_{2v}	C_{2v}
$n=2$	D_{2d}	D_{2d}	D_{2d}	D_{2d}	C_s
$n=3$	D_3	D_3	C_2	C_2	C_2
$n=4$	S_4	C_2	C_2	C_2	C_2

Table 1.7. CCSD(T)(-F12b), i-TTM, and AMOEBA 2009 many-body decompositions (in kcal/mol) for $Na^+(H_2O)_n$ clusters. The energies are divided in contributions from water-water interactions, and from ion-water interactions.

n	symmetry	k-body term	CCSD(T)(-F12b)	i-TTM	AMOEBA 2009
1	C_{2v}	$2B_{total}$	-24.14	-24.86	-23.43
2	D_{2d}	$2B_{ion\text{-}wat}$	-48.20	-49.57	-46.98
		$2B_{wat\text{-}wat}$	0.82	0.83	0.89
		$2B_{total}$	-47.38	-48.74	-46.09
		$3B_{total}$	1.64	2.53	1.62
3	D_3	$2B_{ion\text{-}wat}$	-71.98	-73.93	-70.20
		$2B_{wat\text{-}wat}$	2.86	2.87	3.23
		$2B_{total}$	-69.12	-71.06	66.97
		$3B_{ion\text{-}wat}$	5.46	8.41	5.37
		$3B_{wat\text{-}wat}$	-0.10	-0.13	-0.09
		$3B_{total}$	5.35	8.28	5.28
		$4B_{total}$	-0.19	-0.34	-0.16
4	C_2	$2B_{ion\text{-}wat}$	-79.89	-82.10	-77.61
		$2B_{wat\text{-}wat}$	-3.83	-3.66	-3.00
		$2B_{total}$	-83.72	-85.76	-80.61
		$3B_{ion\text{-}wat}$	3.14	5.77	2.65
		$3B_{wat\text{-}wat}$	1.01	1.03	1.02
		$3B_{total}$	4.16	6.79	3.67
		$4B_{ion\text{-}wat}$	0.34	0.39	0.55
		$4B_{wat\text{-}wat}$	-0.02	-0.03	-0.03
		$4B_{total}$	0.32	0.36	0.52
		$5B_{total}$	-0.03	-0.03	-0.03

the interaction energies E_n of $Na^+(H_2O)_n$ clusters, with $n = 1-4$, are analyzed in terms of individual contributions derived from the many-body expansion (MBE)

$$E_n = \sum_{i=1}^{N} V^{1B}(i) + \sum_{i<j}^{N} V^{2B}(i,j) + \sum_{i<j<k}^{N} V^{3B}(i,j,k) + \cdots + V^{NB}(1,\ldots,N) \tag{1.5}$$

where $N = n+1$ accounting for all water molecules and the alkali metal ion, V^{1B} are the monomer energies, and the V^{kB} are the k-body interactions defined recursively as

$$V^{kB}(1,\ldots,k) = E_k(1,\ldots,k) - \sum_i V^{1B}(i) - \sum_{i<j} V^{2B}(i,j) - \ldots - \sum_{i<j<\cdots<n-1} V^{(k\text{-}1)B}(i,j,\ldots,(k-1)) \tag{1.6}$$

Each term V^{kB} is further decomposed to separate the corresponding contributions associated with the ion-water, $V^{kB}_{\text{ion-wat}}$, and water-water, $V^{kB}_{\text{wat-wat}}$, interactions. In Table 1.7, the individual terms of the MBE calculated with the i-TTM potential for Na^+-H_2O are compared with the CCSD(T)-F12 reference data as well as with the corresponding values obtained with AMOEBA 2009. (Analogous comparisons for the other $M^+(H_2O)_n$ clusters are reported in the Supplementary Information). Independently of the cluster size, the water-water contributions are quantitatively reproduced by the i-TTM potential, which, as discussed in Ref. 90, provides further evidence for the accuracy and transferability of the MB-pol water potential. Non-negligible differences between CCSD(T)-F12 and i-TTM values are instead noticeable in both 2B and 3B contributions, which, as already mentioned above, indicates that purely classical representations may not be sufficient to achieve chemical accuracy in describing the interaction energy of alkali metal ion-water systems. Similar results are obtained with AMOEBA, which, however, displays larger deviations from the CCSD(T)-F12 reference values for the water-water contributions, especially at the two-body level, as expected from the comparisons between the MB-pol and AMOEBA

water potentials.[90] Following Ref. 106, a systematic analysis of many-body contributions to the interaction energies in alkali metal ion-water clusters at both short and long range will be the subject of a forthcoming publication.

While using an accurate representation of the potential energy surface is a prerequisite for a physically correct description of the molecular system of interest, the optimal balance between accuracy and computational efficiency is critical to actually perform molecular simulations long enough to calculate statistically converged quantities. Single point calculations carried out with an in-house code for i-TTM and TINKER (version 6.3)[153,154] for AMOEBA on a single Intel Xeon E5-2640 processor for a cluster consisting of 277 water molecules, 5 Na^+ and 5 Cl^- ions (corresponding to a 1.0 M concentration) indicate that i-TTM is $\sim$3.5 times slower than AMOEBA. More extensive comparisons of the underlying MB-pol water potential with common water models are reported in Refs. 90 and 161.

1.3.3 Comparison between i-TTM and DFT

In this section, the accuracy of the i-TTM potentials in predicting the energetics of $M^+(H_2O)_n$ clusters with $n = 1-4$ is assessed through comparisons with DFT models that are commonly used in simulations of electrolyte solutions. Fourteen functionals (with and without dispersion corrections) are selected for this comparison, including generalized gradient approximation (GGA) functionals (BLYP,[162,163] BLYP-D3BJ, PBE,[164], PBE-D3BJ) meta-GGA functionals (TPSS,[165] TPSS-D3) hybrid-GGA functionals (PBE0,[166] PBE0-D3BJ, B3LYP,[162,163,167] B3LYP-D3BJ), and range-separated hybrid functionals (LC-ωPBE,[149–151] LC-ωPBE-D3BJ, ωB97X,[168] ωB97XD[169]). The D3 dispersion energy corrections with the original[170] or Becke-Johnson (D3BJ)[171] damping functions were applied accordingly.

The DF-MP2 optimized structures of the $M^+(H_2O)_n$ clusters ($n = 1-4$) from the previous section are used in this analysis. The accuracy of the different density functionals is assessed by comparing the DFT interaction energies, E^{DFT}, with the corresponding CCSD(T)(-F12b) reference values, $E^{\mathrm{CCSD(T)(-F12b)}}$, reported in Table 1.5. The deviations, $\delta E = E^{\mathrm{DFT}} -$

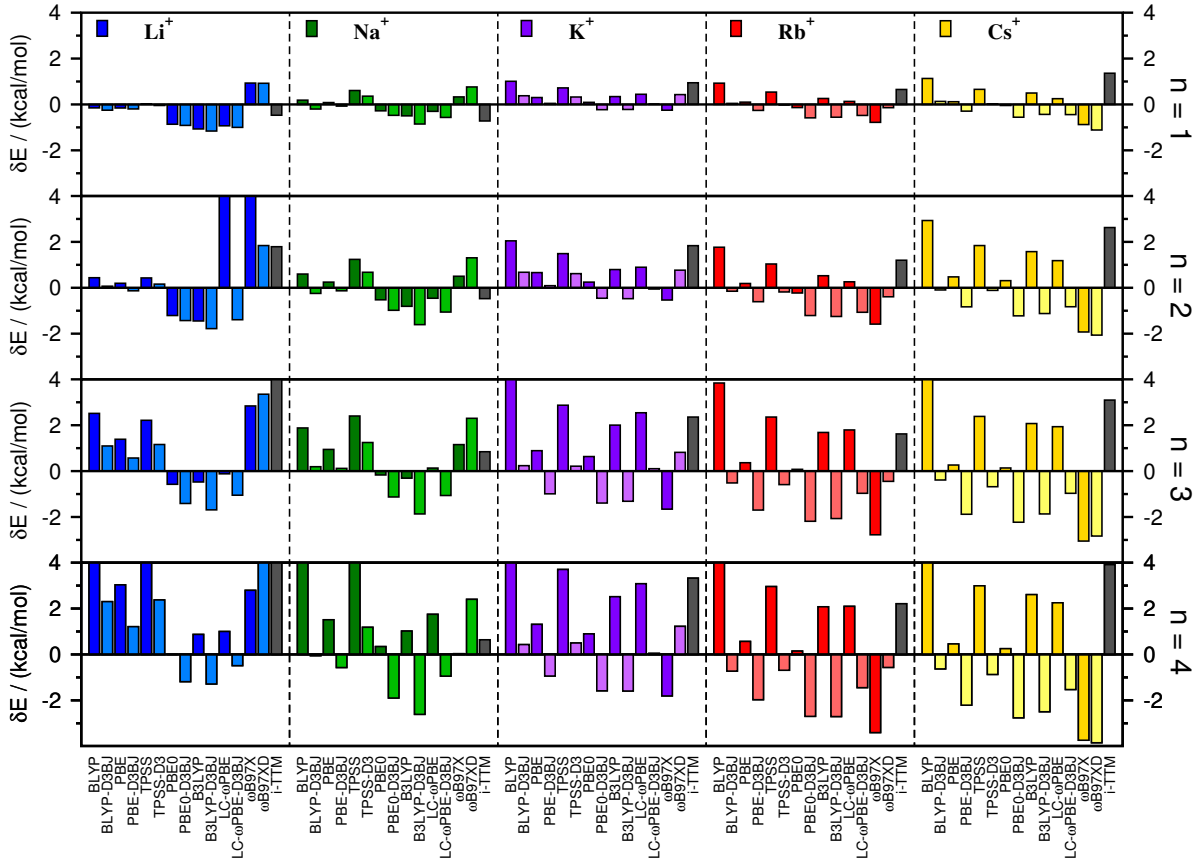

Figure 1.4. Deviations in the interactions energies, $\delta E = E^{\text{model}}$ - $E^{\text{CCSD(T)}(-\text{F12b})}$ (model = DFT or i-TTM) calculated for $M^+(H_2O)_n$ clusters with $n = 1-4$ using different DFT models without (dark color) and with (light color) dispersion energy corrections. Values of $\delta E < -4$ kcal/mol and $\delta E > 4$ kcal/mol are not shown. The dashed lines are guides for the eyes.

$E^{\mathrm{CCSD(T)(-F12b)}}$, are shown in Figure 1.4 as a function of the cluster size. First, this comparison demonstrates that including dispersion corrections does not always improve the accuracy of the DFT results. Second, none of the functionals studied here (with or without dispersion corrections) is able to consistently and accurately describe the interaction energies of $M^+(H_2O)_n$ clusters with different M^+ ions and sizes. For instance, the dispersion corrected GGA PBE-D3BJ and meta-GGA TPSS-D3 are able to describe water clusters of smaller alkali metal ions quite well but display larger errors for the heavier ions (Rb^+ and Cs^+). The hybrid-GGA PBE0 functional shows, overall, the best performance across all clusters studied here, except for the smaller Li^+ clusters. Interestingly, in the case of $Li^+(H_2O)_n$ clusters, all functionals that work best for $n = 1,2$ have significant errors for the larger clusters with $n = 3,4$ and vice versa. Figure 1.4 also shows the i-TTM deviations from the CCSD(T)(-F12b) reference values. Overall, the accuracy of the i-TTM potentials is similar to that of most DFT models, which suggests that they can be used in combination with the halide-water i-TTM potentials of Ref.[105] in molecular dynamics simulations ionic clusters and electrolyte solutions. Work along these lines is ongoing in our group and will be reported in a forthcoming publication.

1.4 Conclusions

To characterize structural, thermodynamic, and dynamical properties, as well as the behavior of electrolyte solutions in different environments, an accurate molecular-level modeling of ion hydration is needed. Following our recent study on the interactions between halide ions and water,[105] we have introduced the i-TTM alkali metal ion-water PESs. The i-TTM energy functions include an explicit treatment of two-body repulsion, electrostatics, and dispersion energy, with all many-body effects being represented by classical induction.

The accuracy of the i-TTM potentials has been assessed through an extensive analysis of both energies and structures of $M^+(H_2O)_n$ clusters, with M^+ = Li^+, Na^+, K^+, Rb^+, and Cs^+, and $n = 1-4$. Comparisons with the corresponding *ab initio* data suggest that the i-TTM

potentials provide an accurate and efficient representation of alkali metal ion-water interactions which correctly include many-body effects. Future efforts will focus on the development of ion-ion potential energy functions compatible with the MB-pol water potential and the i-TTM ion-water potentials, which will enable simulations of electrolyte solutions both in the bulk and at interfaces.

1.5 Conflict of interest

The authors declare no competing financial interests.

1.6 Acknowledgements

We thank Prof. Kirk Peterson for providing the correlation consistent basis sets for potassium, and for advice on the CCSD(T)-F12b calculations. We also want to thank Pushp Bajaj and Shelby Straight for valuable discussions. This research was supported by the National Science Foundation Center for Chemical Innovation "Center for Aerosol Impacts on Climate and the Environment" (grant No. CHE-1305427). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation (grant No. ACI-1053575, allocation TG-CHE110009), and the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract DEAC02-05CH11231.

Chapter 1, in full, is a reprint of the material as it appears in "i-TTM Model for Ab Initio-Based Ion–Water Interaction Potentials. II. Alkali metal ion – water potential energy functions", Riera, M.; Götz, A.W.; Paesani, F.; Phys. Chem. Chem. Phys., 18, 30334, 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 2

Toward Chemical Accuracy in the Description of Ion–Water Interactions Through Many-Body Representations. II. Alkali-Water Dimer Potential Energy Surfaces.

2.1 Introduction

Electrolyte solutions are ubiquitous, occurring in cells and the environment, and mediating many industrial and technological processes.[14–20] In aqueous solutions, the interactions between different ions and surrounding water molecules give rise to specific ion effects, which depend on both the physicochemical properties of the individual ions and the thermodynamic conditions of the solutions.[22] While there have been several experimental and theoretical studies on ion hydration, a definitive picture of how ions are hydrated and to what extent structural differences in the organization of the hydration shells translate into differences in thermodynamic and dynamical properties of electrolyte solutions is still lacking.[23–33]

Significant progress in understanding ion hydration has been made through computer simulations that rely on either force fields[23,59–65] or *ab initio* models.[22,59,111] In particular, a large number of polarizable force fields have been developed to model the behavior of both halide and alkali metal ions in solution.[60,63,64,68,112–119] Most of these models, however, treat the water molecules as rigid, which precludes direct comparisons with experimental vibrational

spectra measured from the gas and the condensed phase.[120–123] A remarkable exception is the AMOEBA force field that includes both intramolecular flexibility and many-body effects through classical polarization.[63] The AMOEBA force field is under continued development, with recent efforts focusing on accurate parameterizations of metal ion-water interactions, including alkali metals,[124] alkaline-earth metals,[125] transition metals,[126,127] and lanthanides and actinides.[128]

Recent years have witnessed the advent of highly-accurate water potential energy functions (PEFs) rigorously derived from the many-body expansion (MBE) of the interaction energies (e.g., CC-pol,[80–82] WHBB,[77,83,84] HBB2-pol,[85,86] and MB-pol[87–89]), enabling computer simulations of water with high accuracy.[90] Among existing many-body PEFs, it has been shown that MB-pol currently provides the most accurate representation of the molecular properties of water from the gas to the condensed phase,[99,100] correctly predicting the vibration-rotation tunneling spectrum of the dimer[87], the relative stability and vibrational spectra of small water clusters[172,173], the structural, thermodynamic, and dynamical properties[89] as well as the IR and Raman spectra of liquid water[129], the vibrational sum-frequency generation (vSFG) spectrum of the air/water interface[174], and the energetics of different ice phases[175].

Building on the MB-pol accuracy for water, a new class of many-body PEFs (MB-nrg for "many-body energy") has recently been introduced to describe halide-water interactions, explicitly including water flexibility.[106] The MB-nrg PEFs are derived entirely from correlated electronic structure data and are expressed through the MBE as[91]

$$E_{\mathrm{N}} = \sum_{i=1}^{N} V^{1\mathrm{B}}(i) + \sum_{i<j}^{N} V^{2\mathrm{B}}(i,j) + \sum_{i<j<k}^{N} V^{3\mathrm{B}}(i,j,k) + \cdots + V^{\mathrm{NB}}(1,\ldots,N) \tag{2.1}$$

where, $V^{1\mathrm{B}}$ is the water monomer energy, and V^{nB} are the n-body (nB) interactions defined

recursively as

$$\begin{aligned} V^{\text{nB}}(1,\ldots,n) = \quad & E_n(1,\ldots,n) - \sum_i V^{\text{1B}}(i) - \sum_{i<j} V^{\text{2B}}(i,j) - \ldots \\ & - \sum_{i<j<\cdots<n-1} V^{\text{(n-1)B}}(i,j,\ldots,(n-1)) \end{aligned} \tag{2.2}$$

Since Eq. 2.1 converges quickly for nonmetallic systems, the MBE provides a rigorous and efficient framework for the development of full-dimensional PEFs in which the low-order interaction terms can be accurately determined using correlated electronic structure methods (e.g., coupled cluster level of theory including single, double, and perturbative triple excitations, CCSD(T), the current "gold standard" for chemical accuracy) and higher-order terms can be effectively represented by many-body induction. In this study, we introduce MB-nrg PEFs for alkali metal ion-water interactions with a specific focus on the 2B $M^+(H_2O)$ term, where M^+= Li^+, Na^+, K^+, Rb^+, and Cs^+. The article is organized as follows: Section 3.2 describes the theoretical and computational details associated with the development of the MB-nrg PEFs. The accuracy of the MB-nrg 2B term is assessed in Section 3 through the analysis of the interaction energies, structures, and harmonic frequencies of $M^+(H_2O)$ dimers. A summary highlighting the main findings and discussing future directions is given in Section 4.

2.2 Theoretical and Computational Methodology

2.2.1 Electronic structure calculations

Depending on the specific alkali metal ion, all 2B energies for the $M^+(H_2O)$ dimers were calculated using either coupled cluster theory with single, double and perturbative triple excitations, CCSD(T), or explicitly correlated CCSD(T)-F12b theory.[130,131] All of the reference energy calculations were performed at the complete basis set (CBS) limit, which was achieved via a two-point extrapolation[132,133] between the values obtained with the correlation-consistent polarized valence triple zeta (aug-cc-pVTZ for H,O, and cc-pwCVTZ for the metal ions) and

quadruple zeta (aug-cc-pVQZ for H,O, and cc-pwCVQZ for the metal ions) basis sets.[134–137] As described in references 133 and 132, the Hartree–Fock energy, the correlation energy and the triples contribution were extrapolated separately for $Li^+(H_2O)$ and $Na^+(H_2O)$, where CCSD(T)-F12b. For $K^+(H_2O)$, $Rb^+(H_2O)$, and $Cs^+(H_2O)$, the Hartree–Fock energy and the correlation energy were extrapolated separately. The reference frequency calculations were obtained with CCSD(T)(-F12b) in conjunction with the aug-cc-pVTZ basis set for H and O, and the cc-pwCVTZ basis set for the metal ions. The reference optimized geometries and energies were obtained with CCSD(T)(-F12b), using aug-cc-pVTZ for H and O, cc-pwCVTZ for the metal ions, and a gradient convergence of 10^{-6} a.u., with a subsequent single point calculation to obtain the CCSD(T)/CBS interaction energy of this structure.

Effective core potentials (ECPs) were used for K^+ (ECP10MDF), Rb^+ (ECP28MDF), and Cs^+ (ECP46MDF).[176] The dipole polarizability of each ion was computed at the CCSD(T) level of theory using the cc-pwCV5Z basis set following the methodology described in Ref. 143. The polarizability values obtained for Li^+, Na^+, K^+, Rb^+, and Cs^+ are 0.0285 $Å^3$, 0.1476 $Å^3$, 0.8184 $Å^3$, 1.3614 $Å^3$, and 2.3660 $Å^3$, respectively, in agreement with other theoretical estimates previously reported in the literature.[144,145] All CCSD(T), CCSD(T)-F12b, and second order Møller-Plesset perturbation theory (MP2) calculations of interaction energies and harmonic frequencies were performed with MOLPRO, version 2015.1.[140] Density functional theory (DFT) calculations of interaction energies and harmonic vibrational frequencies were carried out with Gaussian 09[152] and the aug-cc-pVQZ basis set for all functionals except revPBE, revPBE0, and ωB97M-V[177], for which Q-Chem 5.0[178] was used. The Q-Chem calculations were performed with the aug-cc-pVQZ basis set and an integration grid consisting of 99 radial shells with 590 angular points per shell, (99,590).

2.2.2 Force field calculations

$M^+(H_2O)$ dimer interaction energies and harmonic vibrational frequencies were computed using previously reported classical polarizable forcefields for comparison. The AMOEBA

calculations were performed with version 6.3 of the TINKER package,[153,154] using the AMOEBA 2009 force field, which includes the 2003 parameterizations for both water[179] and alkali metal ions.[124] The TTM-nrg (for "Thole-type model energy") calculations were performed using in-house software. The original TTM-nrg parameters reported in Ref. 107 were used for Li^+ and Na^+. However, since higher accuracy CCSD(T)/CBS reference data for K^+, Rb^+, and Cs^+ were obtained as part of this study using the cc-pwCVTZ and cc-pwCVQZ basis sets, improved TTM-nrg parameters were derived for these ions and used in all comparisons.

2.2.3 MB-nrg functional form

As discussed in Ref. 106, all water properties (i.e., monomer distortion, dipole moment, and polarizability, as well as water-water intermolecular interactions) are described by the MB-pol potential,[87–89] since it has been shown to accurately reproduce structural, thermodynamic, dynamical, and spectroscopic properties of water from the gas to the condensed phase.[99,100,129] Following MB-pol, the interactions between individual ions and water are described through the MBE in Eq. 2.1 and include explicit 2B (ion-water) and 3B (ion-water-water) terms, with all higher-order interactions being implicitly taken into account through many-body induction. The 2B term of the MB-nrg potentials, which is the focus of the present study, includes three contributions:

$$V^{2B}(H_2O-M^+) = V^{2B}_{short}(H_2O-M^+) + V^{2B}_{TTM}(H_2O-M^+) + V^{2B}_{disp}(H_2O-M^+) \qquad (2.3)$$

The electrostatic term (V^{2B}_{TTM}) is a modified version of the extended Thole-type model (TTM) introduced in the TTM4-F water potential,[102] in which the CCSD(T) dipole polarizabilities of the alkali metal ions reported in Section 2.2.1 are used to describe the ion-water induction energy.

The dispersion energy $V_{\rm disp}^{\rm 2B}$ is represented as

$$\begin{aligned} V_{\rm disp}^{2B}(\mathrm{H_2O-M^+}) = &- f_6(D_6^{\rm HM}, R_{\rm H_1M}) C_6^{\rm HM} \frac{1}{R_{\rm H_1M}^6} \\ &- f_6(D_6^{\rm HM}, R_{\rm H_2M}) C_6^{\rm HM} \frac{1}{R_{\rm H_2M}^6} - f_6(D_6^{\rm OM}, R_{\rm OM}) C_6^{\rm OM} \frac{1}{R_{\rm OM}^6} \end{aligned} \tag{2.4}$$

where

$$f_n(D_6R) = 1 - \exp(-D_6R) \sum_{k=0}^{n} \frac{D_6R}{k!} \tag{2.5}$$

is the Tang-Toennies damping function.[158] The C_6 dispersion coefficients were calculated using the exchange-hole dipole model (XDM) developed by Becke and Johnson.[146] The XDM analysis was performed with the Postg software,[147,148] using the wavefunction file obtained from Gaussian 09[152] calculations performed with the LC-ωPBE[149–151] functional and the aug-cc-pV5Z basis set for H and O, and the cc-pwCV5Z basis set for all M^+ ions.

In the MB-nrg energy expression, the classical electrostatic ($V_{\rm TTM}^{\rm 2B}$) and dispersion ($V_{\rm disp}^{\rm 2B}$) contributions are supplemented by a short-range term ($V_{\rm short}^{\rm 2B}$) that effectively recovers quantum-mechanical effects associated with the overlap of the monomer electron densities (e.g., charge transfer, charge penetration, and Pauli repulsion). $V_{\rm short}^{\rm 2B}$ is described by a permutationally invariant polynomial[104] that is smoothly switched to zero for O–M^+ distances larger than a predefined cutoff value

$$V_{\rm short}^{\rm 2B}(\mathrm{H_2O-M^+}) = s_2\left(\frac{R_{\rm OM}-R_i}{R_o-R_i}\right) V_{\rm poly} \tag{2.6}$$

Here, $s_2(x)$ is a switching function defined as

$$s_2(x) = \begin{cases} 1 & \text{if } x < 0 \\ 1 + x^2(2x-3) & \text{if } 0 \le x < 1 \\ 0 & \text{if } 1 \le x \end{cases} \tag{2.7}$$

Table 2.1. Distances and the corresponding variables entering the short range part of the potential, V_{poly}.

d_1	H1	H2	$\xi_1 = e^{-k_{\text{HH}_{\text{intra}}}\left(d_1 - d_{\text{HH}_{\text{intra}}}\right)}$
d_2	O	H1	$\xi_2 = e^{-k_{\text{OH}_{\text{intra}}}\left(d_2 - d_{\text{OH}_{\text{intra}}}\right)}$
d_3	O	H2	$\xi_3 = e^{-k_{\text{OH}_{\text{intra}}}\left(d_3 - d_{\text{OH}_{\text{intra}}}\right)}$
d_4	M	H1	$\xi_4 = e^{-k_{\text{MH}_{\text{coul}}}\left(d_4 - d_{\text{MH}_{\text{coul}}}\right)}/d_4$
d_5	M	H2	$\xi_5 = e^{-k_{\text{MH}_{\text{coul}}}\left(d_5 - d_{\text{MH}_{\text{coul}}}\right)}/d_5$
d_6	M	O	$\xi_6 = e^{-k_{\text{MO}_{\text{coul}}}\left(d_6 - d_{\text{MO}_{\text{coul}}}\right)}/d_6$
d_7	M	L1	$\xi_7 = e^{-k_{\text{ML}}(d_7 - d_{\text{ML}})}$
d_8	M	L2	$\xi_8 = e^{-k_{\text{ML}}(d_8 - d_{\text{ML}})}$

The inner and outer cutoff radii (R_{i} and R_{o}, respectively) of Eq. 2.6 were chosen based on the one-dimensional scans of each $M^+(H_2O)$ dimer. R_{i} corresponds to the O–M^+ distance at which the total and electrostatic energy differ by 0.01 kcal/mol or less, and $R_{\text{o}} = R_{\text{i}} + 1$ Å. Following these criteria, the cutoffs were set to 5.0 Å and 6.0 Å for Li^+, 5.5 Å and 6.5 Å for Na^+, and 6.0 Å and 7.0 Å for K^+, Rb^+ and Cs^+. The permutationally invariant polynomial (V_{poly}) is a function of all distances that involve physical atoms H, O and M^+, as well as the oxygen lone-pair sites (L1 and L2) defined in MB-pol.[87] All distances $d_{\text{n}=1-8}$ entering the expression of V_{poly} are listed in Table 2.1 along with the corresponding variables.

V_{poly} is a function of two different types of variables, intramolecular (ξ_1, ξ_2, ξ_3) and intermolecular Coulomb-type ($\xi_4, \xi_5, \xi_6, \xi_7, \xi_8$) variables, with the permutational invariance imposed with respect to permutations of equivalent sites within the same water molecule (i.e. H_1, H_2). A total of 429 symmetrized monomials form V_{poly}: 3 first-degree monomials formed from all intermolecular variables ($\xi_4, \xi_5, \xi_6, \xi_7, \xi_8$), 15 symmetrized second-degree monomials with at most a linear dependence on intramolecular variables, 49 symmetrized third-degree monomials, 119 symmetrized fourth-degree monomials, 243 symmetrized fifth-degree monomials. V_{poly} thus contains 429 linear fitting parameters (c_i), and 10 nonlinear fitting parameters, $k_{\text{HH}_{\text{intra}}}$, $k_{\text{OH}_{\text{intra}}}$, $k_{\text{XH}_{\text{coul}}}$, $k_{\text{XO}_{\text{coul}}}$, k_{XL}, $d_{\text{HH}_{\text{intra}}}$, $d_{\text{OH}_{\text{intra}}}$, $d_{\text{MH}_{\text{coul}}}$, $d_{\text{MO}_{\text{coul}}}$, d_{ML}. The nonlinear parameters $\gamma_{\parallel}$ and $\gamma_{\perp}$ were set to the same values as in MB-pol, while the dispersion damping coefficients D_6^{OX} and

D_6^{HX} were set at the same values as in the TTM-nrg models.[107] This ensures that the underlying classical representation of alkali metal ion-water interactions is the same in both TTM-nrg and MB-nrg PEFs, with the only differences between the two models being in the description of short-range interactions.

2.2.4 Training sets

The final training sets for the MB-nrg PEFs were generated from a uniform spherical grid of ion positions around a water molecule, including ion-water separations between 1.6 Å to 8 Å . For each position of the ion, different configurations of the water molecule were included. The distortions were obtained by uniformly elongating or shortening one or both O–H bonds, together with a change in the angle, including distortions up to 60 kcal/mol. The original grid, containing approximately 2.7 million configurations, was screened using the RMSD of the distances between all atoms and the center of mass of the $M^+(H_2O)$ dimer as similarity criteria. If the RMSD between two dimers was smaller than a threshold value (TV), the two configurations were considered equivalent and the second one was removed. In the screening process, the specific value of TV varied as a function of the distance between the ion and the oxygen atom of the water molecule, according to

$$TV = \frac{1.0}{1.0 + e^{0.6(k_r - d)}} + 0.005 \tag{2.8}$$

where

$$d = \min(d_{\mathrm{O_1M_1}}, d_{\mathrm{O_2M_2}}) \tag{2.9}$$

and

$$k_r = \begin{cases} k_1 & \text{if} \quad \min(\text{IE}_1, \text{IE}_2) \leq \frac{E_{\min}}{2} \\ k_2 & \text{if} \quad \min(\text{IE}_1, \text{IE}_2) \leq \frac{E_{\min}}{4} \\ k_3 & \text{if} \quad \min(\text{IE}_1, \text{IE}_2) \leq 0.0 \\ k_4 & \text{if} \quad \min(\text{IE}_1, \text{IE}_2) \leq E_{\min} + \Delta E \\ k_5 & \text{if} \quad \min(\text{IE}_1, \text{IE}_2) > E_{\min} + \Delta E \end{cases} \tag{2.10}$$

Here, IE_i is the 2B energy of the ith $M^+(H_2O)$ dimer calculated with the corresponding TTM-nrg PEF[107], and $E_{\min}$ is the 2B energy of the minimum energy configuration $M^+(H_2O)$ dimer. After careful investigation, k_1, k_2, k_3, k_4, and k_5 were set to 10, 9, 8, 7, and 6 for $Li^+(H_2O)$; 11.5, 10.5, 9.5, 8, and 7 for $Na^+(H_2O)$; 12, 11, 10, 9, and 8 for $K^+(H_2O)$; 13, 12, 11, 10, and 8 for $Rb^+(H_2O)$; and 13.5, 12.5, 11.5, 10.5, and 9 for $Cs^+(H_2O)$, respectively, to obtain training sets of roughly $15,000$ configurations for each alkali metal ion-water dimer. ΔE was set to the same value as in the TTM-nrg potentials.[107] These initial grids were further refined to give a more uniform distribution of dimer configurations at both short and long range.

2.2.5 Fitting procedure

Adopting the same fitting protocol established for MB-pol[87,88] and the halide-water MB-nrg PEFs,[106] the linear and nonlinear parameters of the MB-nrg PEFs were optimized with singular value decomposition and the simplex algorithm, respectively, through the minimization of the regularized weighted sum of squared deviations calculated for the corresponding training set (S), commonly known as Tikhonov regularization or ridge regression,[180]

$$\chi^2 = \sum_{n \in S} w_n [V_2^{\text{model}}(n) - V_2^{\text{ref}}(n)]^2 + \Gamma^2 \sum_{i=1}^{429} c_i^2 \tag{2.11}$$

where, the weights, w_n, were set to emphasize dimers with lower total energy according to

$$w(E_i) = \left\{ \frac{\Delta E}{E_i - E_{\min} + \Delta E} \right\}^2 \tag{2.12}$$

Here, ΔE is a parameter used to favorably weight low-energy configurations, and was chosen such that configurations with $E_i > 20$ kcal/mol have weights $w(E_i) \leq 0.25$. Specifically, ΔE was set to 50 kcal/mol for $Li^+(H_2O)$, 40 kcal/mol for $Na^+(H_2O)$, $K^+(H_2O)$, and $Rb^+(H_2O)$, and 30 kcal/mol for $Cs^+(H_2O)$. The regularization parameter, Γ, was set to 5×10^{-4} in order to reduce the variation of the linear parameters without spoiling the overall accuracy of the fits. Figure 2.1 shows the correlation plots for the test sets, which consist of roughly 400 dimer configurations with 2B energies between -40 to 50 kcal/mol. For all five dimers, the MB-nrg PEFs accurately reproduce the CCSD(T) reference energies. The RMSDs of the test sets are 0.23 kcal/mol, 0.09 kcal/mol, 0.14 kcal/mol, 0.05 kcal/mol, and 0.05 kcal/mol for $Li^+(H_2O)$, $Na^+(H_2O)$, $K^+(H_2O)$, $Rb^+(H_2O)$, and $Cs^+(H_2O)$, respectively.

2.3 Results

2.3.1 Validation of the MB-nrg PEFs

The MB-nrg minimum energy structures of each $M^+(H_2O)$ dimer are shown in Figure 2.2, and the associated 2B energies (i.e., interaction energies) are compared in Table 2.2 with the corresponding values calculated using both polarizable force fields and correlated electronic structure methods. While the MB-nrg PEFs quantitatively reproduce the CCSD(T) reference data, providing higher accuracy than MP2, both TTM-nrg and AMOEBA significantly underestimate (up to ∼1.5 kcal/mol) the interaction energies of all dimers.

To assess the accuracy of the MB-nrg PEFs for molecular configurations far from the corresponding minimum-energy structures, the interaction energies calculated along one-dimensional radial scans for different orientations of the M^+ ions relative to a water molecule lying on the *xy*-plane are shown in Figure 2.3. In these calculations, the water molecule was kept in the vibrationally averaged geometry,[142] with the oxygen atom defining the origin of the coordinate frame. The ion position is defined by the distance R from O, and the polar (θ) and azimuthal (ϕ) angles. Although none of these configurations were included in the training sets,

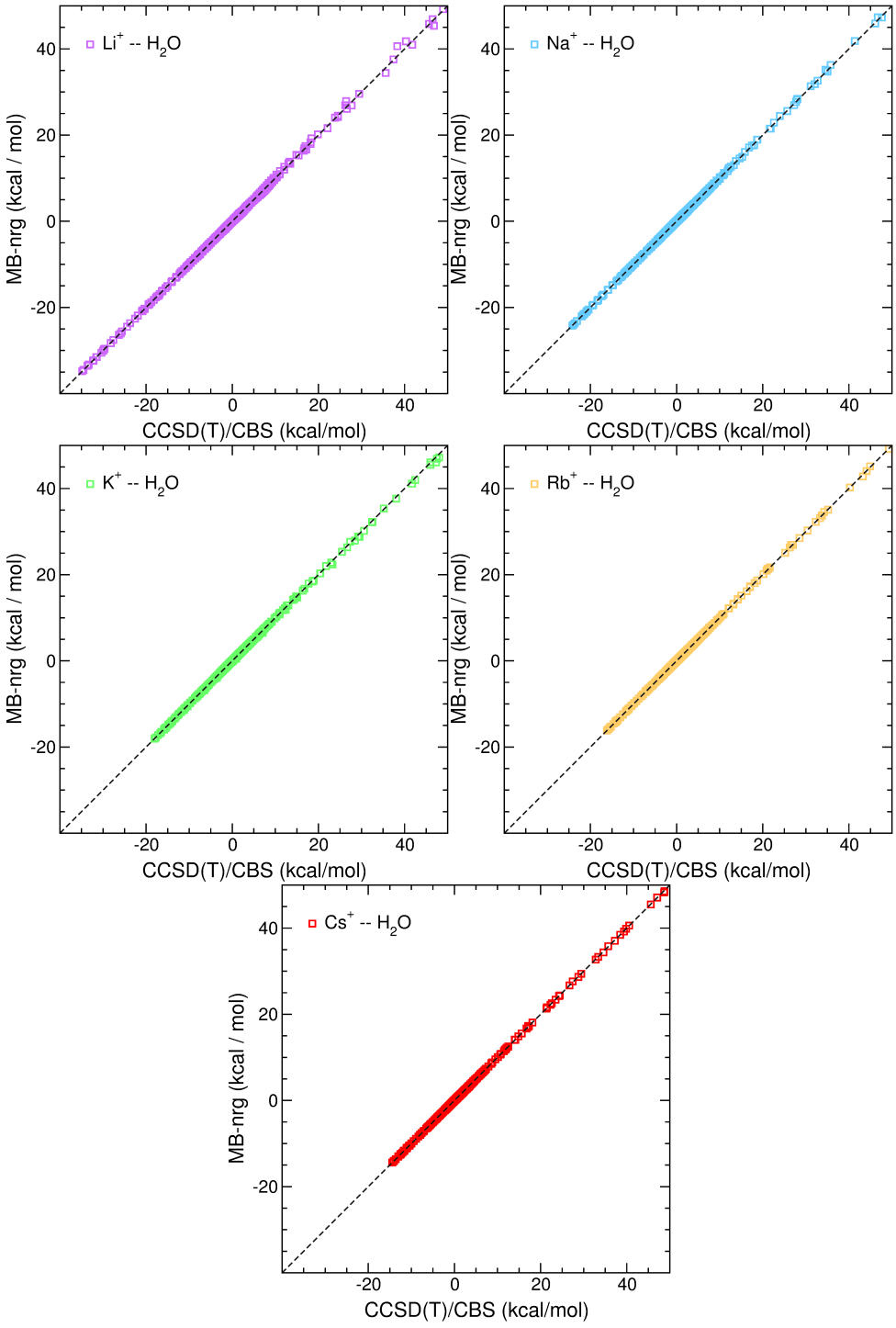

Figure 2.1. Correlation plots of the test set for the 2B energy for $M^+(H_2O)$ ($M^+ = Li^+$, Na^+, K^+, Rb^+, Cs^+) dimers. The CCSD(T)-F12b/CBS 2B energies are plotted on the *x*-axis, and the corresponding 2B energies calculated with the MB-nrg potentials are plotted on the *y*-axis. Each one of the five test sets contains roughly 400 configurations.

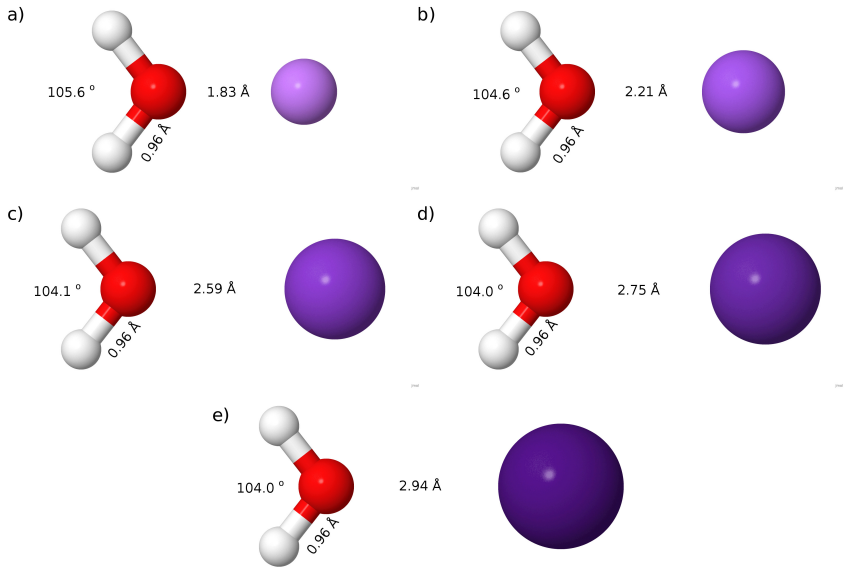

Figure 2.2. Minimum energy geometries for the $M^+(H_2O)$ dimers optimized with the MB-nrg potentials. a) $Li^+(H_2O)$, b) $Na^+(H_2O)$, c) $K^+(H_2O)$, d) $Rb^+(H_2O)$, and e) $Cs^+(H_2O)$.

Table 2.2. Comparison between the 2B energies (in kcal/mol) for the optimized structures at the CCSD(T)(-F12b) level of theory for all $M^+(H_2O)$ dimers, calculated using the MB-nrg potentials and the corresponding values obtained at the CCSD(T)(-F12b) and MP2 levels of theory, as well as using the AMOEBA and TTM-nrg force fields.

Method/Model	$Li^+(H_2O)$	$Na^+(H_2O)$	$K^+(H_2O)$	$Rb^+(H_2O)$	$Cs^+(H_2O)$
AMOEBA[63]	-33.32	-23.43	-17.65	-13.95	-12.64
TTM-nrg[107]	-35.29	-24.91	-17.01	-14.84	-12.75
MP2	-34.51	-23.99	-18.09	-16.31	-14.51
CCSD(T)-F12/CBS	-34.80	-24.18	-18.03	-16.17	-14.43
MB-nrg	-34.79	-24.20	-18.03	-16.18	-14.42

the MB-nrg PEFs (solid lines) correctly reproduce the CCSD(T)(-F12)/CBS values (dotted lines) at all ion-water separations and orientations for all $M^+(H_2O)$ dimers.

An additional test for the accuracy of the MB-nrg PEFs is represented by the analysis of the potential energy angular profiles (PEAPs) along ϕ shown in Figure 2.4 for each $M^+(H_2O)$ dimer, which directly probe the ability of different models to describe the anisotropy of the underlying potential energy surfaces. As for the radial scans, the water molecule was kept in the vibrationally averaged geometry and R was set to 1.83 Å, 2.21 Å, 2.59 Å, 2.75 Å, and 2.94 Å for $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+, respectively, which corresponds to the equilibrium distances within each $M^+(H_2O)$ dimer for $\theta = 0^o$. Independent of M^+, the MB-nrg PEFs quantitatively reproduce the CCSD(T)(-F12)/CBS data at all orientations. By contrast, the TTM-nrg PEFs clearly display some deficiencies in predicting the correct energies of the transition states at $\theta \sim 60^o$, with the deviations from the reference data becoming smaller as the ion size increases, i.e., as the ion-water interactions become more classical-like. Similar behavior was exhibited by the AMOEBA force field in Ref. 107, which reinforces the notion that purely classical PEFs are not capable of correctly reproducing short-range interactions where quantum-mechanical effects (e.g., charge transfer, charge penetration, and Pauli repulsion) due to the overlap of the monomer densities are significant.

2.3.2 Comparisons with DFT

A systematic comparison between the MB-nrg PEFs and different DFT models, with and without the D3 empirical dispersion correction,[170] is presented in this section. The analysis includes results for the energetics and harmonic frequencies calculated with GGA functionals (BLYP[162,163], BLYP-D3, PBE[164], PBE-D3, revPBE[181], revPBE-D3), meta-GGA functionals (TPSS[165], TPSS-D3), hybrid GGA functionals (PBE0[166], PBE0-D3, revPBE0[166,181], revPBE0-D3, B3LYP[162,163,167], B3LYP-D3), and range-separated hybrid functionals (LC-ωPBE[149–151], LC-ωPBE-D3, ωB97X[168], ωB97XD[169], ωB97M-V[177]).

For each $M^+(H_2O)$ dimer, the test set contains $\sim$400 CCSD(T)-F12b dimer configurations

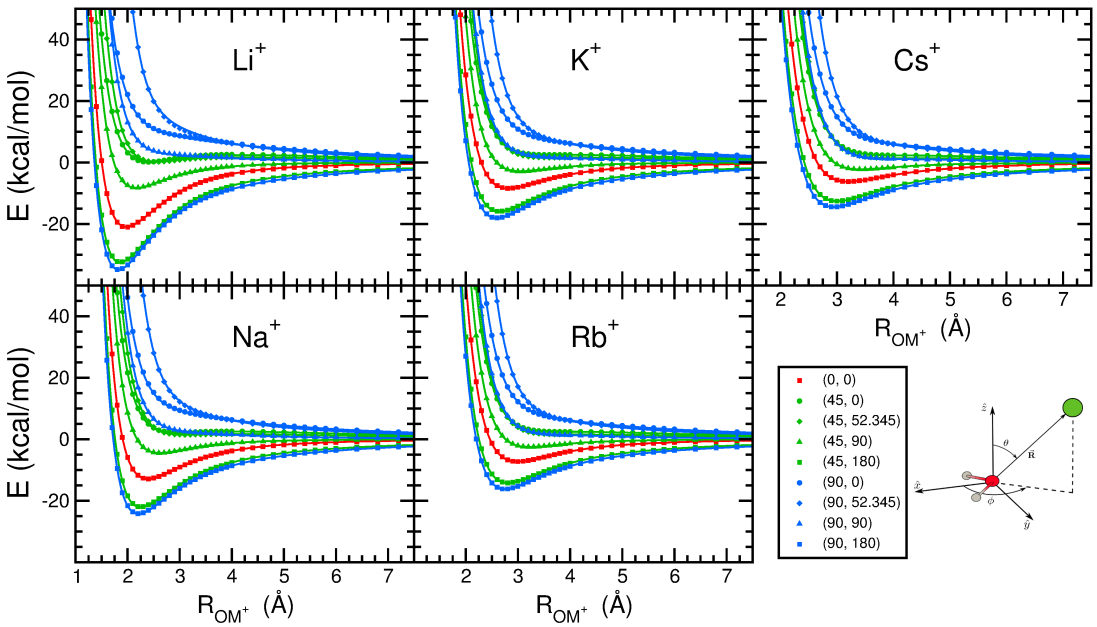

Figure 2.3. Radial scans of the $M^+(H_2O)$ ($M^+ = Li^+$, Na^+, K^+, Rb^+, Cs^+) dimer PESs. The distance between M^+ and O is plotted on the *x*-axis, and the 2B energies are plotted on the *y*-axis. The symbols correspond to the CCSD(T)-F12b/CBS values, while MB-nrg energies are shown with solid lines. The different orientations, θ, ϕ, of M^+ relative to H_2O are given within the parentheses, and the bottom right panel shows the reference frame.

with M^+–O distances between 1.0 Å to 8.0 Å and interaction energies up to 50 kcal/mol. For these test sets, the MB-nrg RMSDs for the $M^+(H_2O)$ dimers are 0.231 kcal/mol, 0.093 kcal/mol, 0.143 kcal/mol, 0.051 kcal/mol, and 0.051 kcal/mol, for $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+, respectively. A comparison of the RMSDs associated with the different models is shown in Figure 2.5, while an analysis of the RMDs as a function of the threshold in the interaction energies and basis sets is presented in the Supplementary Material.

As expected, larger RMSD values are associated with GGA functionals than meta-GGA and hybrid functionals. Interestingly, it appears that the addition of a dispersion correction has either a negative or negligible effect on the performance of all functionals, resulting in larger RMSD values in most cases. Among all functionals considered in this study, ωB97M-V performs best on average, and is the only density functional with RMSDs of less than 0.3 kcal/mol for all five dimers. PBE0 performs second best, and affords RMSDs under 0.3 kcal/mol for all of the dimers except $Li^+(H_2O)$. As noted in previous studies , the accuracy of PBE0 deteriorates

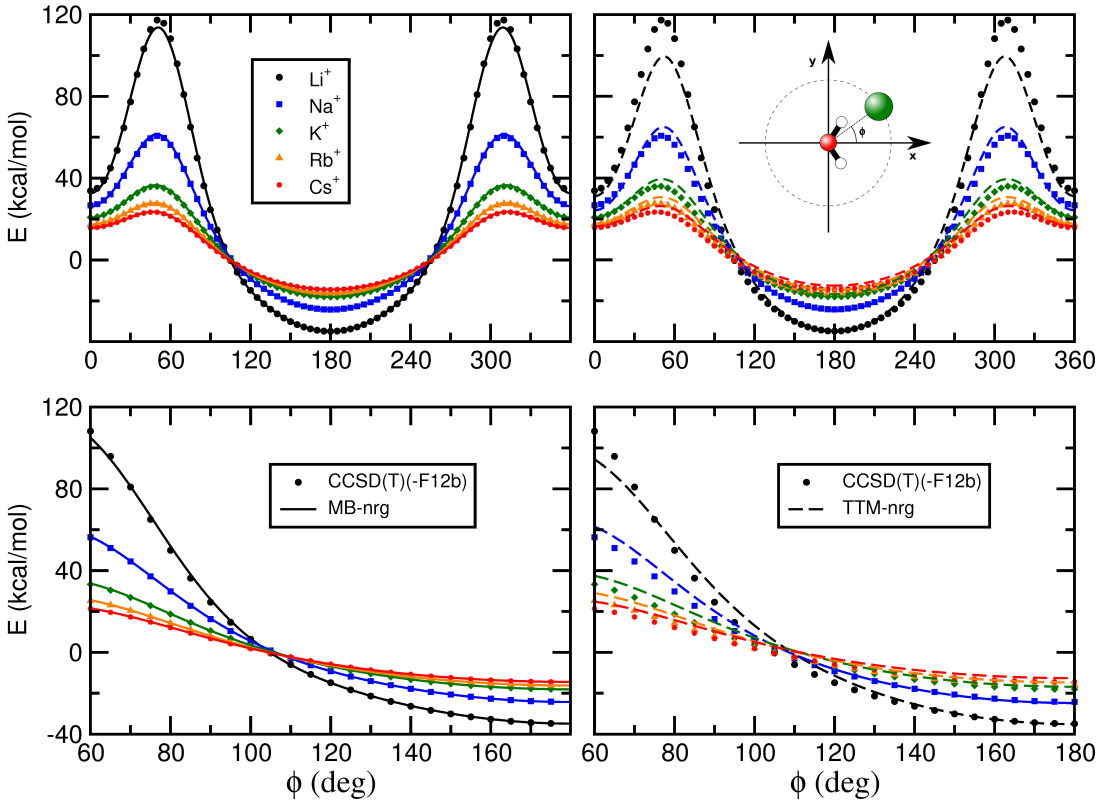

Figure 2.4. Potential energy angular profiles for $M^+(H_2O)$ ($M^+ = Li^+$, Na^+, K^+, Rb^+, Cs^+) calculated as a function of ϕ for $\theta = 90^\circ$, as shown in the top right scheme, and R equal to the minimum energy distance of each dimer, with the water molecule fixed at its vibrationally averaged geometry. The symbols correspond to the CCSD(T)-F12b/CBS values. The MB-nrg energies are shown in the left panels as solid lines, and the corresponding TTM-nrg are shown as dashed lines in the right panels.

(by nearly a factor of two) upon addition of the dispersion correction (e.g., see Ref. 90). The two range-separated functionals without dispersion corrections display opposite behavior for different ions: LC-ωPBE performs well for the smaller cations (Li^+ and Na^+) but not for the larger cations (K^+, Rb^+, and Cs^+), while the reverse is true for ωB97X. It has recently been argued that revPBE and revPBE0 provide a reliable description of aqueous solutions.[182–185] The present analysis of fundamental 2B interactions, however, demonstrates that both functionals actually perform worse than the original PBE and PBE0 functionals for most of the $M^+(H_2O)$ dimers. Finally, the most recent ωB97M-V functional improves by more than a factor of two upon the related ωB97X-D functional for all the cations. For all $M^+(H_2O)$ dimers, the MB-nrg PEFs are always associated with the lowest RMSDs, which, as mentioned above, is effectively uniform and independent of the alkali metal ion.

Two general observations can be made from the analysis of Figure 2.5. First, for any given functional, the RMSD increases upon inclusion of the D3 empirical dispersion correction, and second, the performance of each functional varies in an unpredictable way with the nature of the cation, with the exception of ωB97M-V, which gives uniform RMSD values for the five alkali metal ions.

Direct insights into the overall shape of each $M^+(H_2O)$ PES can be obtained from the analysis of the corresponding vibrational frequencies. Figure 2.6 shows the deviations of the harmonic frequencies of the water bend, and symmetric and asymmetric stretches calculated with both DFT models and MB-nrg PEFs relative to the CCSD(T)(-F12b) reference values. Contrary to what was found for the interaction energies, the inclusion of the D3 dispersion correction has a minimal effect on the vibrational motion of the water molecule. All functionals predict lower harmonic frequencies, with redshifts up to $\sim$60 cm^{-1}, for the bending vibration. Different behavior is instead found for both symmetric and asymmetric stretches, with all GGA and meta-GGA functionals presented in this analysis predicting significantly lower frequencies, with redshifts up to $\sim$150 cm^{-1}, and all hybrid functionals predicting higher harmonic frequencies, with blueshifts up to $\sim$80 cm^{-1}, with the exception of ωB97M-V, which predicts a slightly redshifted frequency

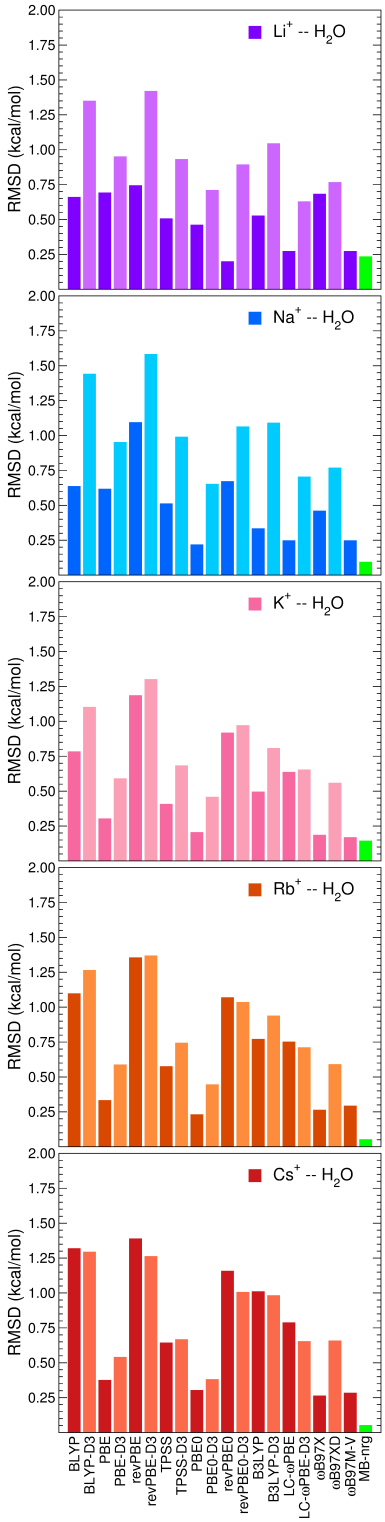

Figure 2.5. RMSDs associated with the 2B energies of the test set with respect to CCSD(T)-F12b/CBS for various DFT models and the MB-nrg potentials.

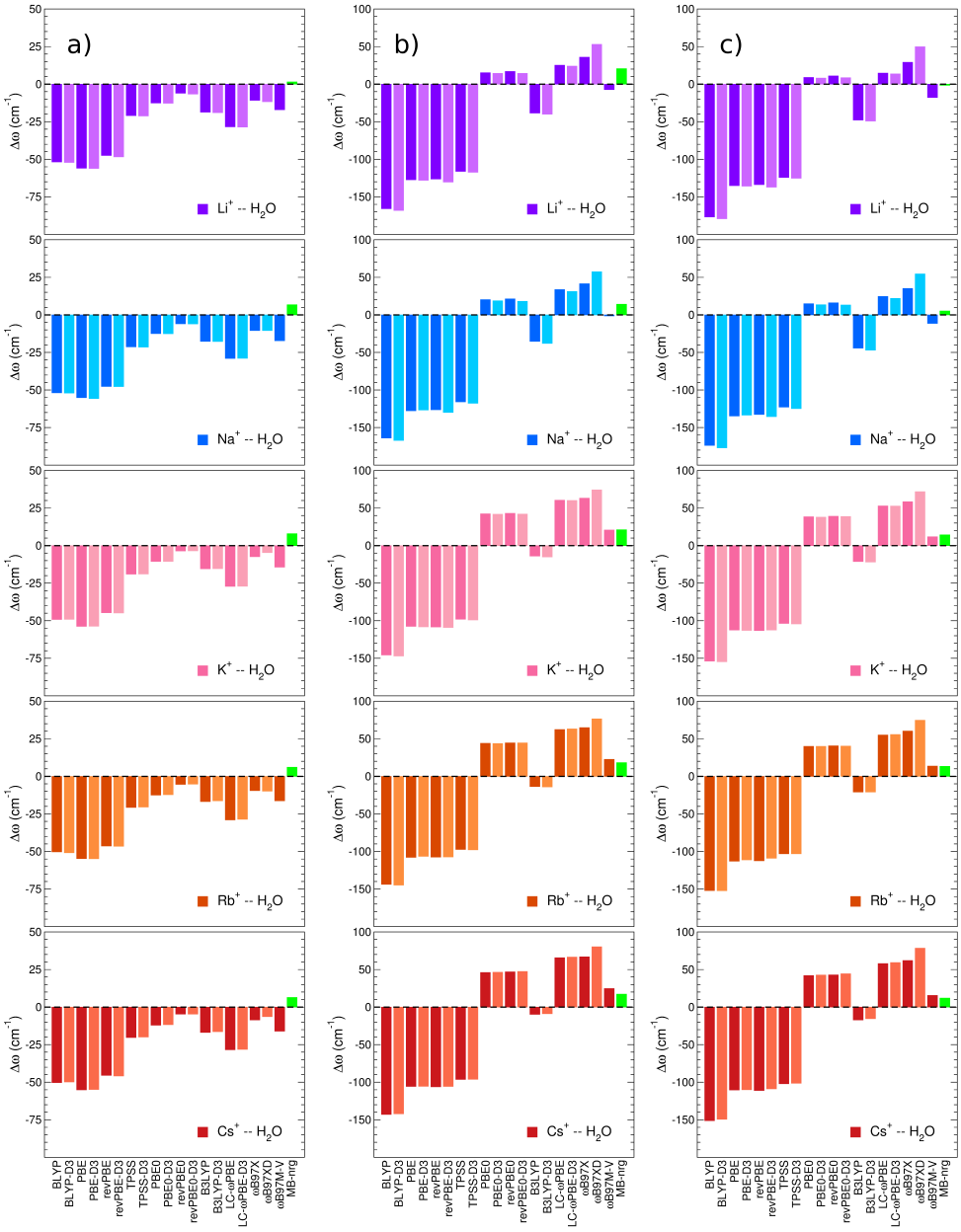

Figure 2.6. Histograms showing the deviations in the a) bending, b) symmetric stretch, and c) asymmetric stretch of the water molecule, in the optimized geometry of each $M^+(H_2O)$ dimer (M^+ = Li^+, Na^+, K^+, Rb^+, Cs^+), using different density functional methods and the MB-nrg potentials from the CCSD(T)(-F12b) normal modes.

for the symmetric and asymmetric stretches in the $Li^+(H_2O)$ and $Na^+(H_2O)$ dimers. The primary exception to this trend is the hybrid B3LYP-D3 functional that underestimates both harmonic frequencies by $\sim$20 cm^{-1}. While revPBE0, B3LYP, and PBE0 provide a consistently reasonable description of the vibrational frequencies of the water molecule in the five $M^+(H_2O)$ dimers, ωB97M-V appears to provide, overall, the closest agreement with the CCSD(T)(-F12b) reference values among all functionals considered in this study, independent of the cation being studied. Interestingly, while ωB97X-D is one of the better functionals for reproducing the harmonic frequency of the bending vibrations, it significantly overestimates both stretching vibrations.

As for the interaction energies, the analysis of the harmonic frequencies in Figure 2.6 demonstrates that the performance of most functionals is particularly sensitive to the nature of the alkali metal ion. On the other hand, independent of the identity of M^+, the MB-nrg PEFs accurately reproduce the coupled cluster reference values, with the deviations never being larger than 10 cm^{-1} for the bending and 20 cm^{-1} for the stretching vibrations.

2.4 Conclusions

In this study, we have introduced the two-body term (V^{2B}) of the MB-nrg PEFs describing the interactions between alkali metal ions and water. By construction, V^{2B} represents the full-dimensional potential energy surface of each $M^+(H_2O)$ dimer, where $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+. Building upon our previous studies on many-body effects in water[87–89] and halide-water systems,[106] V^{2B} is derived entirely from high-level electronic structure data and represented by combining a permutationally invariant polynomial at short range with an underlying classical description of many-body electrostatic and two-body dispersion interactions at all intermolecular separations. As described in Ref. 106, all water-water interactions in the MB-nrg PEFs are represented by the many-body MB-pol PEF that accurately predicts structural, thermodynamic, dynamical, and spectroscopic properties of water from the gas to the condensed phase.[90,100]

The accuracy of the 2B term of alkali metal ion-water MB-nrg PEFs has been assessed

through extensive analysis of both interaction energies and harmonic frequencies for all five $M^+(H_2O)$ dimers. Comparisons with different *ab initio* methods (i.e., coupled cluster, MP2, and several DFT models) and polarizable force fields (i.e., AMOEBA and TTM-nrg) demonstrate that the MB-nrg PEFs provide, in all cases, the highest accuracy, quantitatively reproducing the corresponding coupled cluster reference values. This analysis also shows that the addition of dispersion corrections to popular DFT models tends to have a negative effect on the energetics, while leaving the harmonic frequencies effectively unchanged. Although the magnitude of the DFT deviations from the reference values varies depending on the $M^+(H_2O)$ dimer, the range-separated hybrid ωB97M-V functional overall displays the best performance among all the functionals considered in this study for both interaction energies and harmonic frequencies. On the other hand, both the AMOEBA and TTM-nrg polarizable force fields exhibit noticeable limitations, especially at short-range, as they are not able to correctly represent quantum-mechanical effects (e.g., charge transfer and penetration, and Pauli repulsion) arising from the overlap of the monomer electron densities.

2.5 Supplementary Material

New TTM-nrg parameterization using CCSD(T)/CBS with correlation consistent basis sets as reference. Frequency values for optimized H_2O molecules using all functionals described in Section 2.3.2. Frequency values for optimized $M^+(H_2O)$ dimers, where $M^+ = Li^+$, Na^+, K^+, Rb^+, Cs^+, using all functionals described in Section 2.3.2. Configurations of the optimized dimers using all the above methods.

2.6 Acknowledgements

We thank Dr. Grant Hill and Dr. Kirk Peterson for providing us with the unpublished correlation consistent basis sets for potassium, rubidium and cesium. We also thank Dr. Sandra Brown and Shelby Straight for valuable discussions. This research was supported by the

National Science Foundation Center for Chemical Innovation “Center for Aerosol Impacts on Climate and the Environment” (grant No. CHE-1305427). All expensive calculations for the training set generation (around 300,000 SUs) were performed using resources provided by the Open Science Grid,[??] which is supported by the National Science Foundation and the U.S. Department of Energy’s Office of Science. We thank Edgar Fajardo for his help and technical support on the use of the software in the grid, and the Physics Computing Facilities of UCSD for granting us access to the grid. The DFT calculations were run in the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation (grant No. ACI-1053575, allocation TG-CHE110009) as well as at the Triton Shared Computing Cluster (TSCC) at the San Diego Supercomputer Center.

Chapter 2, in full, is a reprint of the material as it appears in “Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces”, Riera, M.; Mardirossian, N.; Bajaj, P.; Götz, A.W.; Paesani, F.; J. Chem. Phys., 147, 161715, 2017. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Isomeric Equilibria, Nuclear Quantum Effects, and Vibrational Spectra of $M^+(H_2O)_{n=1-3}$ Clusters, with M = Li, Na, K, Rb, and Cs, Through Many-Body Representations

3.1 Introduction

Determining the driving forces and molecular mechanisms that regulate ion hydration is key to the microscopic understanding of fundamental processes that take place in aqueous clusters, solutions, and interfaces. For example, charged species are often found as intermediates in chemical reactions and catalytic processes.[186,187] In biochemistry, ions play a central role in the stabilization of biomolecules[188–190] as well as in mediating protein-protein interactions,[191,192] intracellular signal transduction,[3,4] and enzyme and nucleic acid catalysis.[193–196] Ionic clusters carry electric currents and are involved in the formation and evolution of aerosol particles in the atmosphere,[5,6] while ionic solutions are central to many devices such as electrolytic cells, capacitors, and batteries.[7]

The stabilization of individual ions in solution results from the interplay of energetic contributions associated with ion-solvent and ion-ion interactions, and entropic contributions associated with solvent reorganization required to accommodate the charged species. Given the

central role played by electrolyte solutions in different fields of science and engineering, it is not surprising that much experimental and theoretical work has been and continues to be devoted to the development of a molecular-level understanding of ion hydration.[18,22,59,197] Although the presence of ions is expected to induce changes in the properties of the water hydrogen-bond (HB) network, the precise determination of the extent to which these changes affect structural, thermodynamic, and dynamical properties of aqueous clusters and solutions remains elusive.[198]

A quantitative description of ion hydration is tightly connected to the ability of accurately representing both water-water and ion-water interactions. In this context, computer simulations that rely either on (empirical) force fields (FFs)[23,59–65] or *ab initio*[22,59,111] models based on density functional theory (DFT) provide a promising route for investigating structural, thermodynamic, and dynamical properties of hydrated ions. Several FFs describing the interactions of various ions with water have been reported in the literature and used, with different degrees of success, in molecular dynamics (MD) simulations of ions in water.[60,63,64,68,112–119] Most of these FFs treat the water molecules as rigid, which precludes comparisons with experimental vibrational spectra that directly probe the local hydration structure of ions from small gas-phase clusters to solutions.[120–123] On the other hand, while DFT models should, in principle, provide a parameter-free description of all molecular properties, intrinsic limitations in existing exchange-correlation functionals effectively prevent from achieving chemical and spectroscopic accuracy in *ab initio* MD simulations of ions in water.[199]

The prospect for realistic computer simulations of aqueous systems has gained renewed hope[90] with the advent of accurate potential energy functions (PEFs) for water, such as CC-pol,[80–82] WHBB,[77,83,84] HBB2-pol,[85,86] and MB-pol,[87–89] rigorously derived from the many-body expansion (MBE) of the underlying interaction energies.[91] Not relying on empirical parameterizations, these many-body PEFs allow for actual "predictions" rather than "reproductions" of experimentally measurable quantities.[100] Among existing many-body PEFs, MB-pol has been shown to correctly predict the vibration-rotation tunneling spectrum of the water dimer,[200] the energetics, quantum equilibria, and infrared spectra of small clusters,[88,172,173,201]

the structural, thermodynamic, and dynamical properties of liquid water,[89,99] the energetics of different ice phases,[175] infrared and Raman spectra of liquid water,[129,202,203] the sum-frequency generation spectrum of the air/water interface at ambient conditions,[174] the infrared and Raman spectra of ice I_h.[204] More recently, molecular configurations extracted from classical and path-integral molecular dynamics simulations with MB-pol have been used to determine the electronic band gap of liquid water, both in the bulk and at the air/water interface, through many-body perturbation theory calculations[205]

Building upon the accuracy of MB-pol for water, many-body PEFs (called MB-nrg for "many-body energy") have recently been introduced to describe halide-water and alkali metal ion-water interactions.[106,206] Derived entirely from electronic structure data obtained at the coupled cluster level with single, double, and perturbative triple excitations, i.e., CCSD(T), the current gold standard for chemical accuracy, these MB-nrg PEFs have been shown to outperform both more advanced, polarizable FFs and existing DFT models in the description of the lower-order, two-body (2B) contributions to the corresponding interaction energies.[106,206] Higher order ion-water contributions, at all molecular separations, are taken into account through classical many-body polarization, but not explicitly corrected to reproduce CCSD(T) reference data. Subtle higher-order electronic quantum-mechanical effects (e.g., charge transfer and penetration, and exchange-repulsion), especially at the three-body (3B) level, may be important as suggested in Ref. 207 based on a spectroscopic analysis of small $Na^+(H_2O)_n$ clusters. It should be noted that, compared to the PEF employed in Ref. 207 in which ion-water and water-water interactions are truncated at the 2B and 3B levels, respectively, the MB-nrg PEFs used in the present study include all many-body contributions to both ion-water and water-water interactions. When employed in full-dimensional quantum calculations for $X^-(H_2O)$ and $X^-(D_2O)$ dimers, with X = F, Cl, Br, and I, the MB-nrg PEFs predict vibrational spectra in close agreement with the available experimental data, correctly reproducing anharmonic, nuclear quantum effects, and tunneling splittings.[208]

Along the path connecting ion-water dimers to electrolyte solutions, ion-water clusters

in the gas phase play an important role for understanding ion hydration since, due to their relatively small sizes, they are still amenable to high-level electronic structure calculations while, at the same time, they can be studied experimentally using high-resolution vibrational spectroscopy.[209,210] Continuing our systematic efforts aimed at developing an accurate, molecular-level description of hydration phenomena in different environments and under different thermodynamic conditions, in this study, we investigate the structure, and the temperature-dependent isomeric equilibria and vibrational spectra of $M^+(H_2O)_n$ clusters with M = Li, Na, K, Rb, and Cs, and n = 1-3, at both classical and quantum levels using the MB-nrg PEFs of Ref. 206. The hydration of alkali metal ions has been extensively studied both experimentally[211] and theoretically.[212–214] However, there is not yet a clear consensus between theory and experiments on the hydration structure of these ions in solution.[215] Particularly relevant to our analysis are the infrared photodissociation spectra reported in Refs. 120,122,216–218 and the theoretical studies of Refs. 215,219–222.

3.2 Computatiosnal Details

Following the same theoretical/computational approach used for neutral water clusters,[87,201,223] thermodynamic equilibria between different isomers of $M^+(H_2O)_n$ clusters, with n = 2 and 3, and M = Li, Na, K, Rb, and Cs, were characterized at both classical and quantum levels using replica exchange molecular dynamics (REMD) and replica exchange path-integral molecular dynamics (REPIMD), respectively. Both REMD and REPIMD simulations were carried out with the MB-nrg PEFs introduced in Ref. 206 using 64 replicas between 10 K and 200 K, except for $Li^+(H_2O)_2$ for which the temperatures ranged from 10 K to 250 K, and for $Li^+(H_2O)_3$ and $Na^+(H_2O)_3$ for which the temperatures ranged from 10 K to 350 K. In both REMD and REPIMD simulations, the replicas were distributed according to a geometric temperature progression, which helps ensure efficient exchange between different replicas at both low and high temperature.

In the REPIMD simulations, which provide numerically exact, quantum equilibrium distributions of the different isomers of each $M^+(H_2O)_n$ cluster, each atom was represented by a Feynman's ring polymer with P = 64 beads.[73] After 1 ns of equilibration, classical and quantum equilibrium distributions of the different isomers were obtained from 9 ns of REMD and 3 ns of REPIMD simulations, respectively, with a timestep of 0.2 fs. Instantaneous configurations for the $M^+(H_2O)_n$ clusters were extracted every 2 ps from each replica along the REMD simulations and quenched to identify possible isomers through geometry optimizations carried out by combining linear search and conjugate gradient methods, with a threshold of 10^{-8} kcal mol^{-1}Å^{-1} for the gradients.[224]

To characterize how the presence of different alkali metal ions affect the spatial arrangements and HB network of the solvating water molecules, (anharmonic) infrared spectra of all isomers were calculated by combining the local-mode[75,76] and local-monomer[77] (LM) methods as described in Ref. 201 and briefly summarized here. First, the Hessian matrix for the optimized structure of each isomer was calculated from finite differentiation of the analytic gradients of the MB-nrg PEFs. The local modes of each cluster were obtained using a frequency window of 500 cm^{-1} and a distance-based (Boys) localization criterion. Among the highest $3N$ frequencies, where N is the number of water molecules in the cluster, three local modes were assigned to each water molecule. The intensity of each mode was obtained by computing the dipole surface using the MB-nrg PEFs for the whole cluster, after being centered at the center of mass and oriented along the principal axis. For reference, harmonic spectra were also calculated to quantify the importance of nuclear quantum effects and anharmonicity (see Supporting Information). Harmonic frequencies obtained for both symmetric and asymmetric stretches of the water molecules were found to be blueshifted by $\sim$190 cm^{-1} and 150 cm^{-1} relative to the corresponding LM values, respectively. These differences suggest that some caution should be used in comparing vibrational spectra obtained from classical MD simulations of ions in aqueous clusters and solutions with the corresponding experimental data, since, neglecting nuclear quantum effects, MD simulations significantly underestimate the anharmonicity of the underlying Born-Oppenheimer potential

energy surface and, therefore, effectively provide harmonic vibrational spectra. Temperature dependent IR spectra of each $M^+(H_2O)_n$ cluster were calculated as weighted combinations of the individual IR spectra of the different isomers, with weights corresponding to the isomer fractions as predicted by the REPIMD simulations. It should be noted that, by construction, these temperature-dependent IR spectra are thus approximations to the actual spectra, neglecting both homogeneous and inhomogeneous broadening, and dynamical effects.

3.3 Results and Discussion

Low-lying isomers of all $M^+(H_2O)_n$ clusters with $n = 1-3$ are shown in Figure 3.1. Each isomer is labeled with the acronym Ml(n), where M = L, N, K, R, and C correspond to Li^+, Na^+, K^+, Rb^+ and Cs^+, respectively, l indicates the rank in increasing binding energy (i.e., $l = 1$ indicates the minimum-energy structure), and n is the number of water molecules in the cluster.

3.3.1 $M^+(H_2O)$ clusters

Given both anisotropy and strength of alkali metal ion - water interactions, only one low-lying isomer was identified for each $M^+(H_2O)$ dimer. Since these dimers have been characterized in Ref. 206 through extensive comparisons of MB-nrg with CCSD(T) and DFT calculations, we will briefly summarize here their main features that will serve as a reference for analyzing the evolution of the hydration properties in larger clusters. Figure 3.1 shows that, independent of the specific ion, all dimers display similar structures, with the ion being located along the HOH bisector but on the opposite side of the hydrogen atoms. The different chemical nature (e.g., size) and electronic properties (e.g., charge density) of each alkali metal ion directly correlate with the variation of the dimer binding energies, which decrease by $\sim$20 kcal/mol from $Li^+(H_2O)$ to $Cs^+(H_2O)$.

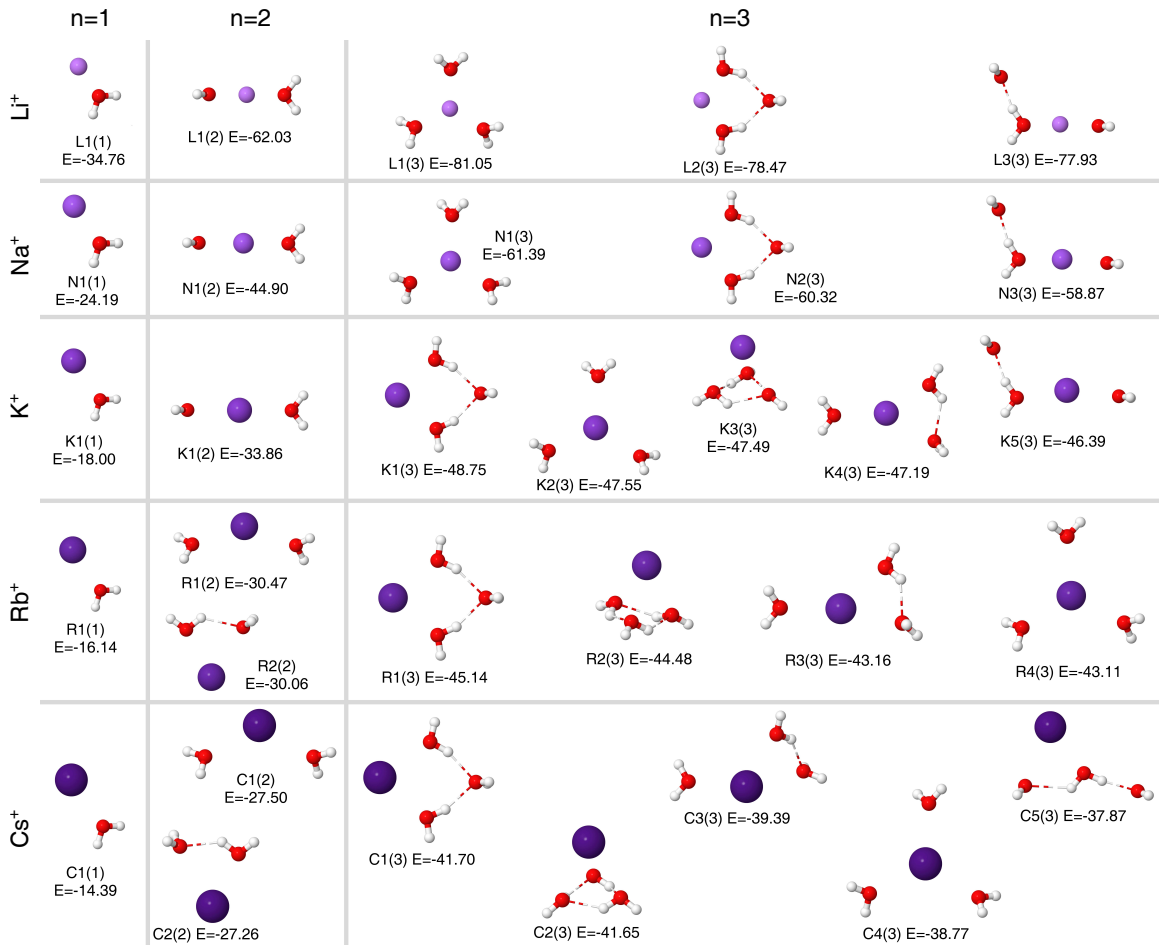

Figure 3.1. Low-lying isomers of all $M^+(H_2O)_n$ clusters with $n = 1-3$. Each isomer is labeled with the acronym Ml(n), where M = L, N, K, R, and C, corresponding to Li^+, Na^+, K^+, Rb^+ and Cs^+, respectively, l indicates the rank in increasing binding energy (i.e., $l = 1$ indicates the minimum-energy structure), and n is the number of water molecules in the cluster. For each isomer, the MB-nrg binding energy (E) is reported in units of kcal/mol.

Table 3.1. LM anharmonic frequencies (in cm^{-1}) calculated with the MB-nrg PEFs for $M^+(H_2O)$ dimers, with M = Li, Na, K, Rb, Cs, compared with the corresponding values calculated for an isolated H_2O molecule. In parentheses are the available experimental values, with the superscripts indicating the reference numbers.

	H_2O	$Li^+(H_2O)$	$Na^+(H_2O)$	$K^+(H_2O)$	$Rb^+(H_2O)$	$Cs^+(H_2O)$
Bend	1582.0 (1595)[225]	1632.6 –	1630.6 –	1625.9 –	1623.3 –	1620.3 –
Symmetric stretch	3656.0 (3657.0)[226]	3628.7 (3631.0)[122]	3642.4 (3634.5)[122]	3636.5 (3636.0)[122]	3636.5 (3637.0)[218]	3634.7 (3635.4)[122]
Asymmetric stretch	3741.4 (3756.0)[226]	3655.8 (3696.0)[122]	3690.6 (3707.0)[122]	3688.2 (3710.0)[122]	3692.0 (3600-3800)[218]	3691.5 (3711.5)[122]

From the analysis of the vibrational frequencies listed in Table 3.1, it is clear that the different strength of the interaction between a water molecule and each of the five alkali metal ions translates into different frequency shifts for the water bending and (symmetric and asymmetric) stretching modes, with Li^+ and Cs^+ providing the largest and smallest shifts, respectively. This can be explained by considering that, due to its higher charge density, Li^+ is more effective in drawing electrons from the water molecule which, in turns, weakens the OH bonds and thus shifts the bending and (symmetric and asymmetric) stretching frequencies more to the blue and the red, respectively, compared to the other M^+ ions. In all cases, the LM vibrational frequencies are in good agreement with the available experimental values[216–218], with deviations ranging from 20 to 40 cm^{-1}, depending on the ion and the mode. These differences may be due to a combination of different factors. First, as mentioned above, the LM method neglects intermonomer, two-mode and higher-order anharmonic couplings, which implies that it only provides an approximate solution to the vibrational Schrödinger equation. In this regard, it should be noted that the LM vibrational frequencies reported in Table 3.1 for a single water molecule differ by $\sim$30 cm^{-1} from the corresponding values obtained from fully coupled quantum calculations in Ref. 101 using the same potential energy surface. Second, the experimental frequencies were obtained from argon-tagging predissociation spectra measured at low, but undetermined, temperature. Besides possible thermal effects that are not considered in the calculations, which strictly correspond to a temperature of 0 K, the effects of Ar-tagging are difficult to quantify. While the interaction between the argon atom and the $M^+(H_2O)$ dimer is supposed to be relatively small for water complexes containing the heavier alkali metal ions, it represents a large fraction of the total interaction energy of the complex with Li^+ and Na^+. Finally, although the MB-nrg PEFs provides an accurate representation of the underlying M^+–H_2O multidimensional potential energy surfaces, they are associated with root-mean-square-deviations (RMSDs) on the order of 0.05 kcal/mol from CCSD(T) reference values, which translates in blueshifts of $\sim$20 cm^{-1} for the water vibrations relative to the corresponding CCSD(T) values.[206]. It should be noted that, for Na^+–H_2O, a similar level of accuracy was reported in Ref. 207 from LM calculations

Table 3.2. Relative ZPE-corrected binding energies (in kcal/mol) calculated for $M^+(H_2O)_2$ clusters, with M = Li, Na, K, Rb, and Cs, using the MB-nrg PEF. The corresponding binding energies on the underlying Born-Oppenheimer potential energy surface are shown in Figure 3.1.

Isomer	$Li^+(H_2O)_2$	$Na^+(H_2O)_2$	$K^+(H_2O)_2$	$Rb^+(H_2O)_2$	$Cs^+(H_2O)_2$
1	0.00	0.00	0.00	0.00	0.89
2	–	–	–	1.37	0.00

performed with a strictly 2B ion-water PEF.

3.3.2 $M^+(H_2O)_2$ clusters

As shown in Figure 3.1, the addition of a second water molecule leads to some diversification among $M^+(H_2O)_2$ clusters with different M^+ ions. Specifically, while Li^+, Na^+, and K^+ form only one low-lying, linear isomer, with the two water molecules being coordinate through the oxygen atom at opposite sides of the ion, a second, cyclic isomer exists for both $Rb^+(H_2O)_2$ and $Cs^+(H_2O)_2$, in which the two water molecules are hydrogen bonded to each other. As a result, the vibrational spectra of $M^+(H_2O)_2$ clusters with M = Li, Na, and K at 0 K show the same features as those of the corresponding dimers, with the symmetric stretch at $\sim$3640 cm^{-1} and the asymmetric stretch at $\sim$3700 cm^{-1}. Since only one low-lying isomer exists for these complexes, the IR spectra are predicted to remain effectively unchanged in the temperature range from 0 K to 200 K examined in this study. In addition, due to the absence of other isomers, only REMD simulations were carried out for these clusters.

As shown in Figure 3.1, the Born-Oppenheimer energy differences between the two lowest isomers of $Rb^+(H_2O)_2$ and $Cs^+(H_2O)_2$ are 0.41 kcal/mol and 0.24 kcal/mol, respectively. However, after inclusion of the corresponding zero-point energies (ZPEs) within the harmonic approximation, Table 3.2 shows that the C2(2) isomer becomes the actual ground state of $Cs^+(H_2O)_2$, lying 0.89 kcal/mol below C1(2), while R1(2) remains the lowest-lying isomer of $Rb^+(H_2O)_2$, lying 1.37 kcal/mol below R2(2). The different role played by nuclear quantum effects on the relative stability of $Rb^+(H_2O)_2$ and $Cs^+(H_2O)_2$ clusters becomes even more

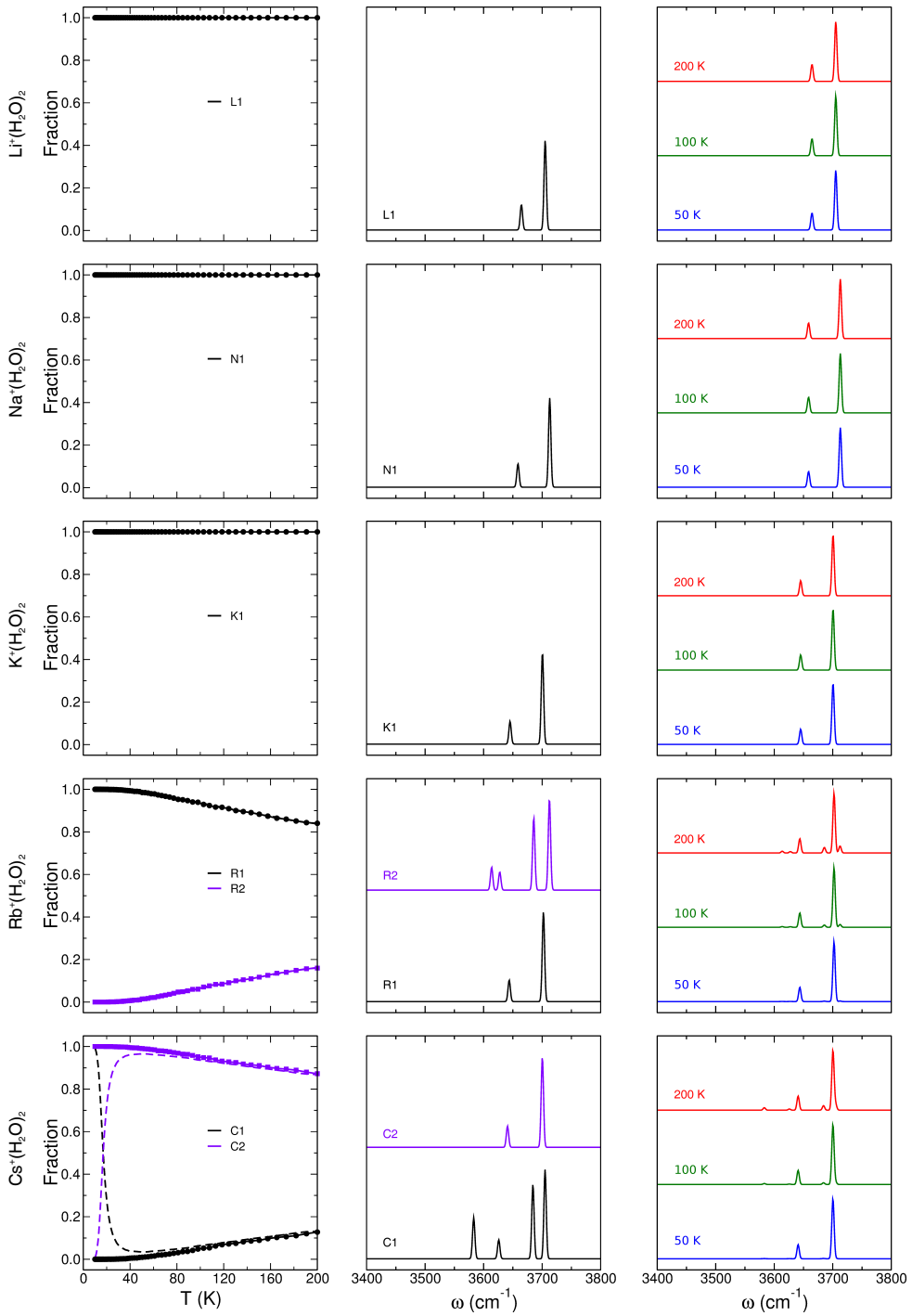

Figure 3.2. Left column: Fractions of different isomers of each $M^+(H_2O)_2$ cluster, with M = Li, Na, K, Rb, and Cs, calculated using REMD (dashed lines) and REPIMD (solid lines) simulations as a function of temperature. Middle column: LM anharmonic vibrational spectra calculated for individual isomers at 0 K and smoothed using Gaussians with standard deviations of 2 cm^{-1}. Right column: Vibrational spectra calculated for each $M^+(H_2O)_2$ cluster at 50 K, 100 K, and 200 K as weighted combinations of the individual isomer spectra (middle column), with weights corresponding to the isomer fractions as predicted by the REPIMD simulations (left column).

apparent from the analysis of the REMD and REPIMD results shown in Figure 3.2. While the fractions of R1(2) and R2(2) calculated at both classical and quantum levels are identical at all temperatures, the REMD simulations predict that C1(2) is the most abundant isomer below 10 K before sharply decreasing up to ∼40 K and then slowing increasing as the temperature increases. In contrast, by correctly including nuclear quantum effects, the REPIMD simulations show that C2(2) is the most stable isomer over the entire temperature range, from 0 K to 200 K.

Since the bent isomer of both $Rb^+(H_2O)_2$ and $Cs^+(H_2O)_2$ clusters effectively dominates at all temperatures, the corresponding IR spectra are very similar to those calculated for the analogous clusters with Li^+, Na^+, and K^+, displaying two main peaks associated with the symmetric and asymmetric stretches of the water molecules. Minor spectral features associated with "cyclic" isomers appear at temperature higher than 100 K. Overall, the calculated spectra are in agreement with the available experimental data.[122,218] However, it should be noted that the experimental spectra also display minor features above 3750 cm^{-1} corresponding to $\Delta K = \pm 1$ rotational subbands of the water asymmetric stretching mode,[216] which, by construction, cannot be reproduced by the LM method. As for $Na^+(H_2O)$, the present calculations for $Na^+(H_2O)_2$ in the ground state are in close agreement with previous theoretical predictions.[207].

3.3.3 $M^+(H_2O)_3$ clusters

The addition of a third water molecule in the hydration shell of M^+ ions gives rise to a more diverse group of low-lying isomers which now includes linear, branched, and cyclic structures, with and without hydrogen-bonded water molecules. As shown in Figure 3.1, there are significant structural differences between the minimum-energy configurations of $M^+(H_2O)_3$ clusters with $M^+ = Li^+$ and Na^+, and those containing Rb^+ and Cs^+, with the former corresponding to symmetric, branched configurations with the ion in the center of a triangle that has the water molecules at the three vertices, and the latter corresponding to rhombic configurations with the three water molecules and the ion located at the four vertices. Another important structural difference between $M^+(H_2O)_3$ clusters containing the lighter (Li^+ and Na^+) and the heavier

Table 3.3. Relative ZPE-corrected binding energies (in kcal/mol) calculated for $M^+(H_2O)_3$ clusters, with M = Li, Na, K, Rb, and Cs, using the MB-nrg PEF. The corresponding binding energies on the underlying Born-Oppenheimer potential energy surface are shown in Figure 3.1.

Isomer	$Li^+(H_2O)_3$	$Na^+(H_2O)_3$	$K^+(H_2O)_3$	$Rb^+(H_2O)_3$	$Cs^+(H_2O)_3$
1	0.00	0.00	0.88	0.15	0.00
2	4.28	2.94	0.00	1.31	0.73
3	3.56	3.18	2.40	0.93	1.21
4	–	–	1.08	0.00	0.64
5	–	–	2.01	–	2.96

(K^+, Rb^+, and Cs^+) ions is the appearance of configurations in which the ion is located on top of planar, hydrogen-bonded triangular structures formed by the three water molecules. Importantly, when ZPE contributions are taken into account within the harmonic approximation, Table 3.3 shows that the relative stability of the $K^+(H_2O)_3$ and $Rb^+(H_2O)_3$ isomers changes, with the branched configuration becoming the ground state as seen for $Li^+(H_2O)_3$ and $Na^+(H_2O)_3$. The rhombic structure, C1(3), remains the ground state for $Cs^+(H_2O)_3$ upon inclusion of the ZPE, although the branched configuration becomes the second low-lying configuration (see Table 3.3). These ion-dependent structural differences clearly point to a subtle balance between ion-water and water-water interactions on the underlying Born-Oppenheimer potential energy surface, and nuclear quantum effects.

As shown in Figure 3.3, while the relative stabilities of the isomers of both $Li^+(H_2O)_3$ and $Na^+(H_2O)_3$ do not change as the temperature increases, with the corresponding minimum-energy structures being the only observed structures below 200 K, the isomeric equilibria for the other $M^+(H_2O)_3$ clusters depend significantly on temperature and display nonnegligible nuclear quantum effects which, depending on M^+, may persist up to temperatures higher than 100 K. For $K^+(H_2O)_3$, isomer K2(3) remains the dominant isomer at the quantum mechanical level, although the fraction of isomer K4(3) increases continuously starting at $\sim$40 K. Importantly, at the classical level, isomer K1(3), corresponding to the minimum-energy structure on the Born-Oppenheimer potential energy surface, disappears quickly between $\sim$10 K and $\sim$40 K, while the

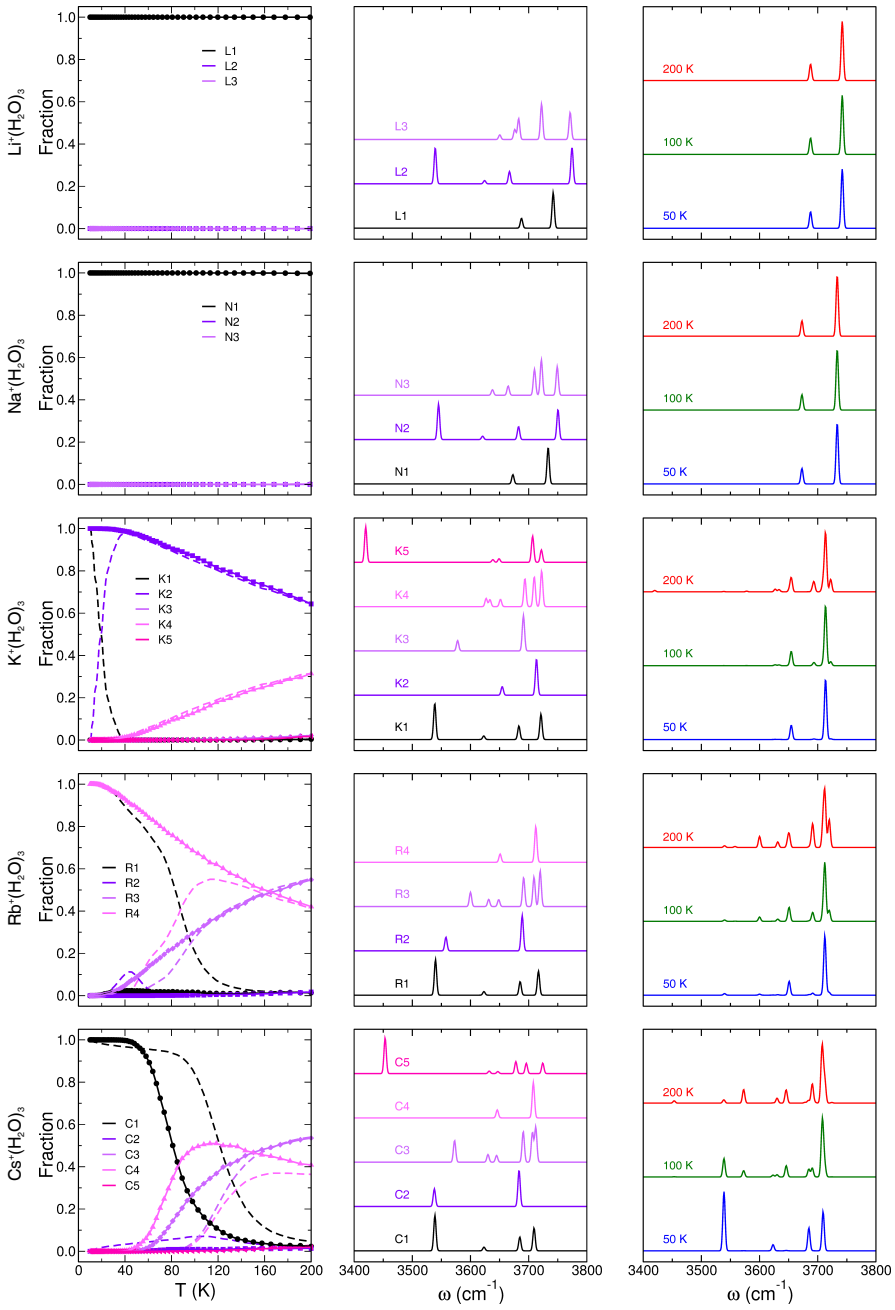

Figure 3.3. Left column: Fractions of different isomers of each $M^+(H_2O)_3$ cluster, with M = Li, Na, K, Rb, and Cs, calculated using REMD (dashed lines) and REPIMD (solid lines) simulations as a function of temperature. Middle column: LM anharmonic vibrational spectra calculated for individual isomers at 0 K and smoothed using Gaussians with standard deviations of 2 cm^{-1}. Right column: Vibrational spectra calculated for each $M^+(H_2O)_3$ cluster at 50 K, 100 K, and 200 K as weighted combinations of the individual isomer spectra (middle column), with weights corresponding to the isomer fractions as predicted by the REPIMD simulations (left column).

fraction of isomer K2(3) steeply increases in the same temperature interval. This inversion in the relative stability of the two lowest-lying isomers can be explained by considering the higher entropy associated with isomer K2(3), in which all OH bonds of the three water molecules are free. It should also be noted that the fractions of the different isomers of $K^+(H_2O)_3$ calculated at both classical and quantum levels become effectively indistinguishable above $\sim$40 K.

Nuclear quantum effects clearly play a major role in determining the relative stability of the different isomers of $Rb^+(H_2O)_3$ and $Cs^+(H_2O)_3$. In particular, the minimum-energy structure, isomer R1(3), of $Rb^+(H_2O)_3$, which is the most abundant isomer at the classical level up to $\sim$80 K, is not present at any temperature when ZPE contributions are taken into account within the REPIMD formalism. Similarly, the fraction of isomer R2(3), which appears between $\sim$20 K and $\sim$40 K in the classical simulations, is negligible over the entire temperature range in the corresponding REPIMD simulations, which instead predict isomers R4(3) and R3(3) to be the most stable structures below and above $\sim$160 K, respectively. The analysis of the REMD and REPIMD results shows that the crossover from quantum to classical behavior in $Rb^+(H_2O)_3$ also takes place at $\sim$160 K. While both REMD and REPIMD simulations predict isomer C1(3) to be the most stable structure of $Cs^+(H_2O)_3$ at low temperature, nuclear quantum effects modify the relative stability of the other isomers. Specifically, REPIMD simulations predict isomers C4(3) and C3(3) to become the most stable structures at $\sim$80 K and $\sim$150 K, respectively, with negligible fractions of isomers C2(3) and C5(3) present above $\sim$100 K. In contrast, at the classical level, isomer C1(3) remains the dominant structure up to $\sim$120 K, after which isomer C3(3) becomes the most stable structure. The REMD simulations also predict isomer C4(3) to represent a relatively large fraction between $\sim$120 K and $\sim$200 K, as well as a small, but nonnegligible, fraction of isomer C2(3) over the entire temperature range, and a slowly increasing fraction of isomer C5(3) above $\sim$100 K. This analysis suggests that, for $Cs^+(H_2O)_3$, the crossover temperature from classical to quantum behavior occurs above 200 K.

Since each $M^+(H_2O)_3$ isomer is characterized by a specific hydrogen-bonding motif, the temperature dependence of the isomeric equilibria is directly mirrored by the evolution of the

corresponding infrared spectra. As follows from L1(3) and N1(3) being the only isomers present in the entire temperature range, the calculated infrared spectra of both $Li^+(H_2O)_3$ and $Na^+(H_2O)_3$ are characterized by two peaks corresponding to the symmetric and antisymmetric stretches of the three equivalent water molecules of the L1(3) and N1(3) isomers, respectively, which are the only isomers present over the entire temperature range. The LM frequencies, $\sim$3680 cm^{-1} and $\sim$3740 cm^{-1} for $Li^+(H_2O)_3$, and $\sim$3660 cm^{-1} and $\sim$3730 cm^{-1} for $Na^+(H_2O)_3$, are within 20-30 cm^{-1} of the experimental frequencies measured in Ar-tagged predissociation experiments.[122]

Richer infrared spectra are calculated at different temperatures for $M^+(H_2O)_3$ clusters with the heavier ions. Since, at the quantum mechanical level, the dominant isomer of $K^+(H_2O)_3$ is isomer K2(3), which is isostructural with L1(3) and N1(3), the calculated infrared spectra are characterized by analogous symmetric and asymmetric stretch peaks at $\sim$3655 cm^{-1} and $\sim$3710 cm^{-1}, with additional spectral features associated primarily with isomer K4(3) appearing at higher temperature. Overall, good agreement is found with the Ar-tagged predissociation measurements reported in Ref. 122, although the experimental spectra also shown a feature at 3546 cm^{-1}, which is an unambiguous signature of a weak hydrogen bond. The LM spectra calculated for the individual isomers of $K^+(H_2O)_3$ indicates that this spectral feature can be associated with isomer K1(3). Since this isomer is predicted not to be present for the untagged clusters under the equilibrium conditions enforced by the REPIMD formalism, it is possible that it was observed in experiments due to stabilization effects associated with Ar-tagging or because it was trapped under nonequilibrium conditions during the supersonic expansion.

As discussed above, nuclear quantum effects are particularly pronounced in $Rb^+(H_2O)_3$, with two isomers, R3(3) and R4(3), effectively being present with significant fractions at all temperatures. This is mirrored by the corresponding infrared spectra calculated as a function of temperature, which are dominated by isomer R4(3) at low temperature and progressively evolve into combined spectra, with dominant features between 3600 cm^{-1} and 3750 cm^{-1} primarily associated with both R3(3) and R4(3) isomers, and smaller peaks at 3540 cm^{-1} and 3550 cm^{-1}

associated with strong hydrogen bonds characteristic of isomers R1(3) and R2(3), respectively. Also for $Rb^+(H_2O)_3$, good agreement is found with the experimental spectra reported in Ref. 227, especially for the frequencies of the water symmetric and asymmetric stretches. However, it should be noted that the experimental intensity of the hydrogen-bond peak at 3550 cm^{-1} is significantly higher than that predicted by the present LM calculations. Such differences between theoretical predictions and experimental measurements, and may be due to the presence of the tagging species in the experiments, inaccuracies in the representation of the many-body dipole moment surface in the calculations, or a combination of the two.

$Cs^+(H_2O)_3$ is the first $M^+(H_2O)_3$ cluster in which the most stable isomer, C1(3), corresponds to a rhombic structure with the ion and the three hydrogen-bonded water molecules at the vertices. As a result, the infrared spectra at low temperature display the characteristic signature of hydrogen bonding at $\sim$3540 cm^{-1}, which progressively disappears as the temperature increases and isomer C3(3) and C4(3) become increasingly more stable. The LM spectra presented in Figure 3.3 can be compared with two sets of measurements reported in Refs. 122 and 227. Both measured and calculated spectra display well-defined features at $\sim$3540 cm^{-1} and $\sim$3710 cm^{-1} corresponding to water hydrogen-bonded and free OH stretches, respectively. The most recent measurements[227] also show a distinct band at $\sim$3650 cm^{-1}, which can be assigned to the water symmetric stretches of isomer C4(3), with a small shoulder at $\sim$3540 cm^{-1}, which correspond to hydrogen-bonded OH stretches of isomer C3(3). While good agreement between measured and calculated spectra is found for all these features, the experimental spectra also show a band at $\sim$3600 cm^{-1}, which is not present in any of the LM spectra calculated for the individual isomers and, consequently, does not show up in the corresponding temperature-dependent spectra. This discrepancy between experiment measurements and theoretical predictions appears to require further investigation. It should also be noted that, also for $Na^+(H_2O)_3$, the LM results presented here are in agreement with those reported in Ref. 207 using a different ion-water and water-water PEFs truncated at the 2B and 3B levels, respectively.

3.4 Conclusions

We have presented a detailed analysis, at both classical and quantum levels, of the temperature dependence of isomeric equilibria and vibrational spectra of $M^+(H_2O)_n$ clusters, with n = 1 - 3, and $M^+ = Li^+$, Na^+, K^+, Rb^+, and Cs^+, calculated with the MB-nrg PEFs introduced in Ref. 206. Nuclear quantum effects are found to play a negligible role in determining the relative stability of different isomers of clusters containing the lighter ions, Li^+ and Na^+, with one structure of each $M^+(H_2O)_n$ cluster effectively dominating the isomeric equilibria over the temperature range between 0 K and 200 K. Although the low number of water molecules considered in this study is not sufficient to fully complete the first solvation shell around Li^+ and Na^+, the most stable structures of $Li^+(H_2O)_n$ and $Na^+(H_2O)_n$ clusters tend to resemble those found in solution, with the ion at the center of the cluster. At the classical level, competing effects between ion-water and water-water interactions become more pronounced for the heavier ions, which results in a larger number of low-lying isomers, some of which appear to serve initial seeds for extended hydrogen-bonded networks. Within this picture, K^+ holds a special place since, at the classical level, the relative stability of its clusters with water follows a similar trend as that of the lighter ions at low temperature and as that of the heavier ions at high temperature.

When nuclear quantum effects are explicitly taken into account within the path-integral formalism, the relative stability of $K^+(H_2O)_n$ clusters is largely modified by zero point energy contributions which shift the isomeric equilibria toward structures with no hydrogen-bonded water molecules, such as those predicted for $Li^+(H_2O)_n$ and $Na^+(H_2O)_n$ clusters. Similar trend is followed by $Rb^+(H_2O)_n$ clusters, which, however, exhibit increasing stability for isomers containing hydrogen-bonded water molecules. Water clusters containing Cs^+ are the first ones for which the most stable structures correspond to isomers where the ion is interacting with a network of two and three hydrogen-bonded water molecules.

For all clusters, good agreement is found between anharmonic vibrational spectra calculated within the local monomer approximation and the available experimental data.[122,216–218]

Besides providing further support for the accuracy of the MB-nrg PEFs, this level of agreement emphasizes the importance of taking properly into account anharmonic and quantum effects in computer simulations for a correct description of $M^+(H_2O)_n$ clusters. Importantly, the possibility to directly connect experimental measurements with computer simulations at the molecular level provides hope for the development of a consistent picture of ion hydration from the gas to the condensed phase.

3.5 Supplementary Material

Geometries for all isomers of each $M^+(H_2O)_n$ cluster, along with the corresponding REMD and REPIMD fractions, and harmonic and anharmonic vibrational spectra calculated at 0 K, along with their corresponding intensities.

3.6 Acknowledgements

We thank Pushp Bajaj and Dr. Andreas Götz for valuable discussions. This research was supported by the National Science Foundation through grant no. CHE-1453204 and the Air Force Office of Scientific Research through grant no. FA9550-16-1-0327, and used resources of the Extreme Science and Engineering Discovery Environment (XSEDE),[?] which is supported by the National Science Foundation through grant no. ACI-1053575, under allocation TG-CHE110009 on Comet at the San Diego Supercomputer Center and Stampede2 at the Texas Advanced Computing Center. M.R.R. was supported by a Software Fellowship from the Molecular Sciences Software Institute, which is funded by the National Science Foundation through grant no. ACI-1547580.

Chapter 3, in full, is a reprint of the material as it appears in "Isomeric Equilibria, Nuclear Quantum Effects, and Vibrational Spectra of $M^+(H_2O)_{n=1-3}$ Clusters, with M = Li, Na, K, Rb, and Cs, through Many-Body Representations", Riera, M.; Brown, S.E.; Paesani, F.; J. Phys. Chem. A, 122, 5811, 2018. The dissertation author was the primary investigator and author of

this paper.

Chapter 4

Many-Body Effects Determine the Local Hydration Structure of Cs^+ in Solution

Determining the driving forces and molecular mechanisms that regulate the hydration properties of alkali-metal ions is fundamental for a microscopic understanding of several processes taking place in aqueous clusters, solutions, and interfaces, which, in turn, have major implications for various fields of science and engineering. For example, lithium salts are widely used in rechargeable batteries,[110,228] and are effective agents in the treatment of manic-depressive illness.[229] Sodium and potassium ions play important roles in the stabilization of biomolecules,[188–190] intracellular signal transduction,[3,4] and enzyme and nucleic acid catalysis.[193–196] On the other hand, the heavier alkali-metal ions (i.e., rubidium and cesium) are not as ubiquitous in either the environment or in living systems. However, the accident at the Fukushima Daiichi reactor in 2011 has drawn attention to the importance of a molecular-level understanding of adsorption and desorption processes of soluble radionuclides, such as ^{137}Cs, for the development of efficient technologies for the treatment of nuclear waste.[230]

In this context, deriving a molecular picture of the hydration properties of Cs^+ ions presents significant challenges to both experiment and theoretical modeling due to the intricate interplay between Cs^+–water and water–water interactions, which, being of similar magnitude, are difficult to disentangle. As a result, large variability is found in the values reported in the literature for both Cs^+ coordination number and spatial extent of the hydration shells

around Cs^+ ions in solution. From large-angle X-ray (LAXS) and double difference infrared (DDIR) spectroscopic experiments performed on a 2.0 M cesium iodide solution, the mean distance between Cs^+ and the oxygen atoms of the water molecules (Cs^+–O) within the first hydration shell was estimated to be 3.07 Å, corresponding to a coordination number of 8.[211] Anomalous X-ray diffraction patterns measured for a 3 molal cesium iodide solution were used to determine a coordination number of 7.9, assuming a Cs^+–O distance of 3.0 Å.[231] Values in the range of 7.7 – 8 for the coordination number and 3.1 Å – 3.2 Å for and Cs^+–O distance were reported from subsequent neutron and X-ray diffraction measurements.[232,233] Various theoretical and computational approaches, including *ab initio* molecular dynamics (AIMD), quantum mechanics/molecular mechanics (QM/MM), and classical molecular dynamics (MD), have also been used to investigate the hydration structure of Cs^+, resulting in predictions for the coordination number and Cs^+–O distance in the range of 7 – 10 and 3.0 – 3.3 Å, respectively.[60,234–238]

As already implicit in the "flickering clusters of hydrogen-bonded molecules" picture of liquid water proposed by Frank and Wen,[239] different hydrogen-bonding arrangements in solution result from the delicate balance of many-body interactions,[90,91,98] which is further modulated by nuclear quantum effects.[240,241] By combining measurements of vibration-rotation tunneling spectra with analogous theoretical calculations and MD simulations, it was argued that short-range many-body effects were likely the most important flaws in commonly used empirical water models.[242–245] The development of explicit many-body potential energy functions (PEFs), rigorously derived from the corresponding many-body expansions of the underlying interaction energies, represents a significant step toward predictive computer simulations of aqueous systems at the molecular level.[77,80–89,105–107,161,206] By combining the MB-pol PEF,[87–89] which correctly predicts the properties of water across different phases,[100] with the TTM-nrg and MB-nrg PEFs,[106,206] which accurately describe molecular interactions between halide and alkali-metal ions with water, theoretical studies have been carried out to characterize the tunneling dynamics in $X^-(H_2O)$ and $X^-(D_2O)$ dimers, with X = F, Cl, Br, and I,[208] as well as quantum isomeric

equilibria in small $M^+(H_2O)_n$ clusters, with M = Li, Na, K, Rb, and Cs.[246]

In this study, a series of molecular models built upon a hierarchy of approximate representations of Cs^+–water interactions is used in MD and path-integral molecular dynamics (PIMD) simulations[73] to investigate many-body effects in the hydration structure of Cs^+ ions. Specifically, the analyses presented in the following include the empirical point-charge force field obtained by combining the TIP4P-Ew model for water[108] with the corresponding Cs^+–water parameterization introduced in Ref. 109, hereafter referred to as TIP4P-Ew, the polarizable TTM-nrg PEF, with an implicit representation of N-body (NB) effects based on classical induction,[107] and the many-body MB-nrg PEF, with either explicit inclusion of two-body (2B) interactions in addition to the same classical NB term adopted by TTM-nrg, hereafter referred to as (2B+NB)-MB-nrg, or explicit inclusion of both two-body (2B) and three-body (3B) interactions in addition to the TTM-nrg classical NB term, hereafter referred to as (2B+3B+NB)-MB-nrg.[206] The accuracy of the four different models in reproducing the hydration structure of Cs^+ in diluted solutions as well as the role played by nuclear quantum effects are then assessed through systematic comparisons between measured and simulated L_1-edge and L_3-edge extended X-ray absorption fine structure (EXAFS) spectra. Specific details about both simulations and experimental measurements are reported in the Supporting Information.

To first assess the ability of the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg models to describe Cs^+–water interactions, Fig. 4.1 shows the correlation plots between 2B and 3B energies calculated with each model and the corresponding reference values obtained at the coupled cluster level of theory with single, double, and iterative triple excitations, CCSD(T). In this analysis, 2B and 3B energies are calculated for 380 distinct $Cs^+(H_2O)$ and $Cs^+(H_2O)_2$ configurations with distorted water geometries, respectively. In the case of the (rigid) TIP4P-Ew model, the CCSD(T) reference energies are calculated for the same dimer and trimer configurations used for the other models after rescaling both OH bonds and HOH angles of the water molecules to match the TIP4P-Ew parameterization. Specific details about the CCSD(T) calculations are reported in the Supporting Information.

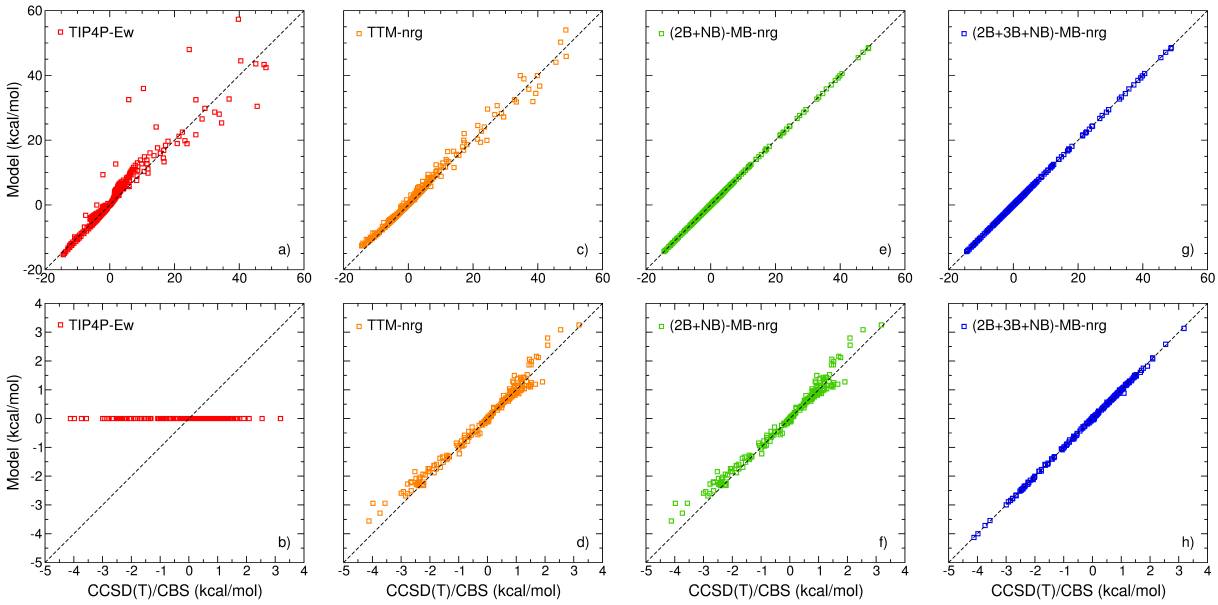

Figure 4.1. Correlation plots for 2B (top panels) and 3B (bottom panels) Cs^+–water interaction energies for the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models. On the x-axes are the CCSD(T) reference data while on the y-axes are the corresponding values calculated with each model.

This large RMSE can be explained by considering that, as in all empirical pairwise-additive force fields, the 2B term of TIP4P-Ew does not strictly represent 2B interactions but effectively also accounts for higher-order interactions that, by construction, are not explicitly included in the model (e.g., 3B interactions in panel b). As shown by the correlation plots for the TTM-nrg model (panels c and d), implicit inclusion of many-body effects through classical polarization significantly improves the agreement with the CCSD(T) data for both 2B and 3B energies, resulting in RMSEs of 1.23 kcal/mol and 0.14 kcal/mol, respectively. Explicit representations of 2B interactions, which is accomplished in the MB-nrg PEFs (panels e and g) by introducing permutationally invariant polynomials (PIPs) that effectively represent non-classical contributions to molecular interactions (e.g., charge transfer and penetration, and Pauli repulsion),[106,206] leads to further reduction of the 2B RMSEs to 0.05 kcal/mol. Finally, the correlation plots shown in panels d, f and h demonstrate that purely classical representations of 3B interactions adopted by the TTM-nrg and (2B+NB)-MB-nrg models are not sufficient to quantitatively reproduce the CCSD(T) data, resulting in RMSEs of 0.17 kcal/mol. Significantly

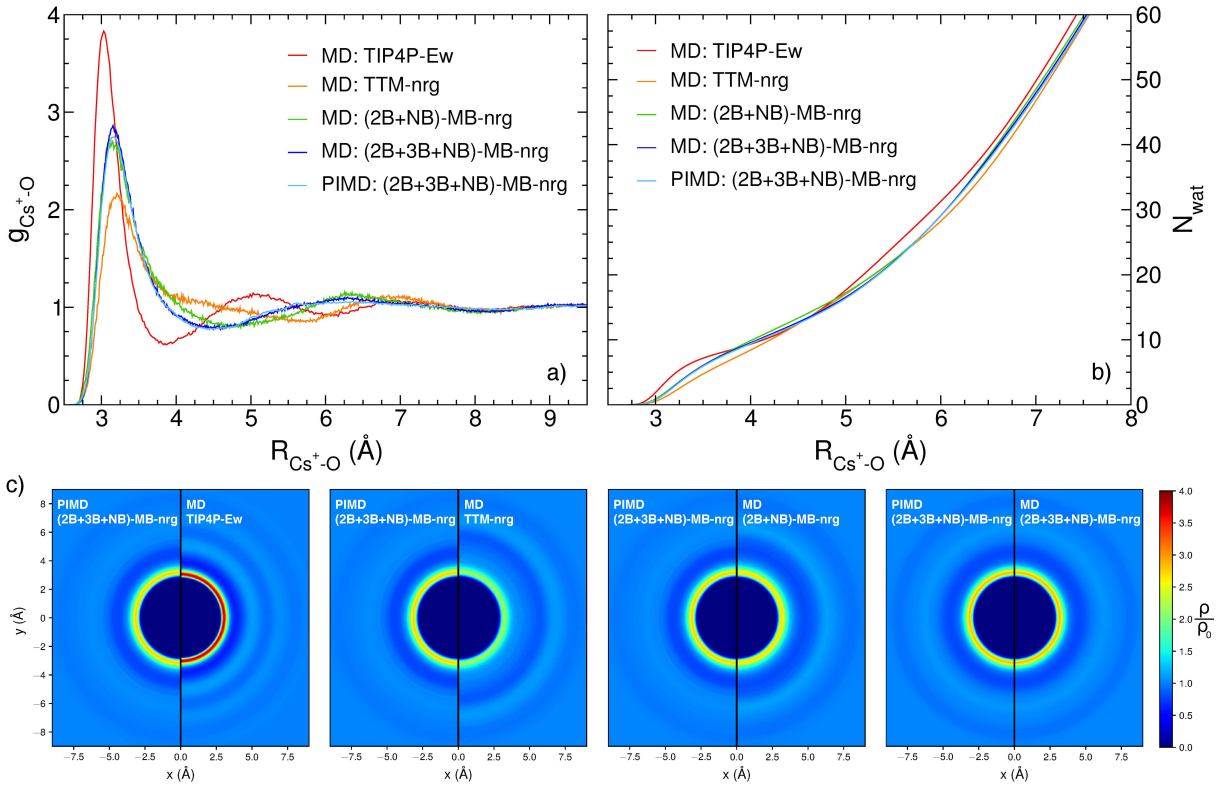

Figure 4.2. a) Radial distribution functions (RDFs) describing the spatial correlation between the Cs^+ ion and the oxygen (O) atoms of the water molecules calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg model. b) Corresponding coordination numbers calculated as a function of the Cs^+–O distance. c) Two-dimensional plots comparing the density profiles calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models on the left of each panel with the corresponding results obtained from PIMD simulations with the (2B+3B+NB)-MB-nrg model.

higher accuracy in the description of 3B energies, with an associated RMSE of 0.02 kcal/mol, is exhibited by the (2B+3B+NB)-MB-nrg model, which supplements the classical description of 3B Cs^+–water interactions adopted by the TTM-nrg and (2B+NB)-MB-nrg models with explicit PIPs representing non-classical 3B contributions.

Having established the accuracy of the four models in reproducing the lower-order, and more relevant, many-body effects in Cs^+–water interactions, Fig. 4.2 analyzes how the differences found in the correlation plots of Fig. 4.1 impact the hydration structure of Cs^+ in solution. The comparison between the Cs^+–O radial distribution functions (RDFs), g_{Cs^+-O},

calculated from MD and PIMD simulations with the four models (Fig. 4.2a) clearly shows that TIP4P-Ew predicts a more structured distribution of water molecules around Cs^+, with a sequence of well-defined peaks located at ∼3.0 Å, ∼5.0 Å, ∼7.0 Å, and ∼9.0 Å. Inclusion of an implicit description of many-body effects through a classical polarization term as implemented in the TTM-nrg model effectively leads to the collapse of the hydration shell located at ∼5.0 Å in the TIP4P-Ew RDF. This is accompanied by the broadening of the first peak, corresponding to an expansion of the 1^{st} hydration shell, whose position consequently shifts to relatively larger Cs^+–O distances. As a result of this structural reorganization of the water molecules around Cs^+, the second and third peaks (i.e., 2^{nd} and 3^{rd} hydration shells) in the TTM-nrg RDF effectively correspond to the third and fourth peaks (i.e., 3^{rd} and 4^{th} hydration shells) in the TIP4P-Ew RDF.

Qualitatively different hydration structures are predicted by the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models, which progressively include explicit representations of 2B and 3B contributions to Cs^+–water interactions. Compared to the TIP4P-Ew RDF, the RDFs calculated with both MB-nrg models display a lower but broader first peak that extends up to ∼4.5 Å, as well as a second, broader peak extending from ∼4.5 Å to ∼8.0 Å. Contrary to TTM-nrg predictions, the MB-nrg RDFs display a well-defined sequence of hydration shells, with a sharper first peak at ∼3.15 Å and a second, broader peak at ∼6.2 Å. Although comparisons between RDFs calculated from simulations with (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg suggests that the inclusion of an explicit 3B term in the (2B+3B+NB)-MB-nrg model only leads to minor changes to the overall hydration structure of Cs^+, it nevertheless contributes to sharpening the interstitial region between the first and second peak, providing further evidence for the importance of non-classical 3B effects in Cs^+–water interactions. Finally, the comparison between the RDFs obtained from MD and PIMD simulations with the (2B+3B+NB)-MB-nrg model indicates that nuclear quantum effects play a minimal role in determining the hydration structure of Cs^+ in solution.

Fig 4.2b shows that the differences in the RDFs obtained from simulations with the four models directly translate into different distributions of water molecules around the Cs^+ ion. In particular, while MD simulations with the TIP4P-Ew model predict that ∼8 molecules are within

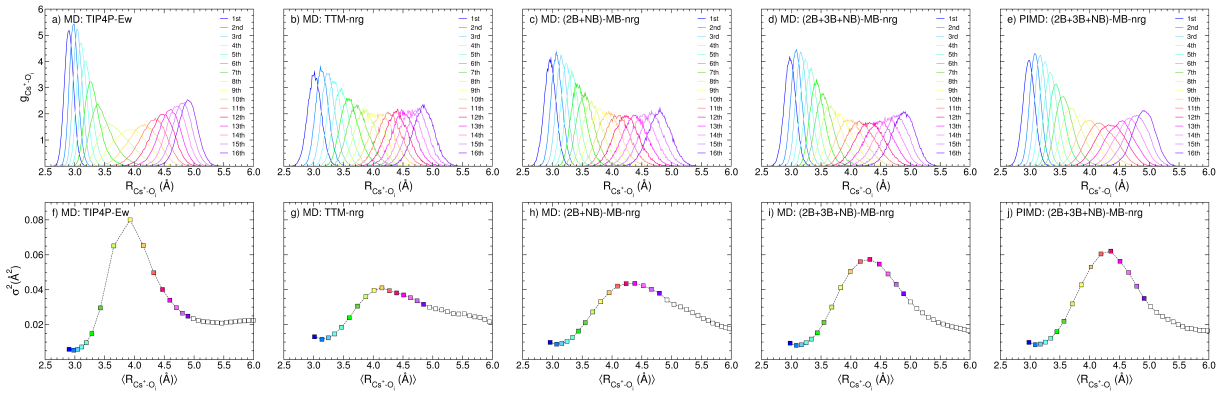

Figure 4.3. a-e) Incremental radial distribution functions (i-RDFs) calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg model. f-j) Average distances (x-axis) and associated variances (y-axis) calculated for each i-RDF shown in the corresponding top panel, using the same color scheme. Values for water molecules with $i > 16$ are shown as empty square symbols.

the 1st hydration shell of Cs^+, the lack of a well-defined 2nd hydration shell makes it difficult to unambiguously determine a coordination number from the corresponding TTM-nrg simulations. The wider 1st hydration shells predicted by simulations with both MB-nrg models result in relatively larger coordination numbers ($\sim$10), with negligible differences due to nuclear quantum effects. The more compact nature of the hydration structure predicted by the TIP4P-Ew model is visually apparent in the two-dimensional density plots shown in the first panel of Fig. 4.2c. Fig. 4.2c also shows that the systematic inclusion of more accurate representations of many-body effects, from the TTM-nrg to the (2B+3B+NB)-MB-nrg model, leads to the progressive reshaping of the hydration shells, particularly at short Cs^+–water distances. Overall, the comparisons in Fig. 4.2 demonstrate that many-body effects have significant impact on the local Cs^+ hydration structure, while pairwise-additive representations, which approximate Cs^+–water interactions in a mean-field sense, become increasingly more accurate at larger Cs^+–water distances ($R_{Cs^+-O} >$ 7.0 Å).

These incremental radial distribution functions (*i*-RDFs) calculated from MD simulations with TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg as well as from PIMD simulations with (2B+3B+NB)-MB-nrg are shown in panels a-e of Fig. 4.3, respectively. While

the TIP4P-Ew *i*-RDFs show a clear separation between water molecules belonging to the 1st and 2nd hydration shells, with the 9th molecule located in the interstitial region between the two shells, all other models predict a more gradual transition between the fist two hydration shells. This transitions is extremely hard to detect in the TTM-nrg *i*-RDFs and becomes more distinct as both many-body and nuclear quantum effects are explicitly taken into account in the PIMD simulations with the (2B+3B+NB)-MB-nrg model.

The different evolution of the hydration shells predicted by the four models becomes more evident from the analysis of the average distances, $\langle R_{Cs^+-O_i} \rangle$, and associated variances, σ_i^2, calculated in Fig. 4.3f-j for the individual *i*-RDFs. MD simulations with the TIP4P-Ew model predict narrower distributions for water molecules closer to the Cs^+ ion ($\langle R_{Cs^+-O_i} \rangle \leq 3.4$ Å), indicating relatively stronger Cs^+–water interactions, as well as the broadest distributions for water molecules (8th to 11th) located at the boundary of the 2nd and 3rd hydration shells, with σ_i^2 becoming effectively constant for the *i*-RDFs corresponding to water molecules with $i \geq 16$. The distinction between water molecules located between 3.5 Å and 4.5 Å and those at both smaller and larger distances from the Cs^+ ion becomes less marked when many-body effects are implicitly included in simulations with the TTM-nrg model. Fig. 4.3 shows that explicit account for many-body Cs^+–water interactions and nuclear quantum effects in simulations with the MB-nrg models progressively leads to sharper distributions of σ_i^2 as a function of $\langle R_{Cs^+-O_i} \rangle$. Although these distributions are qualitatively similar to that obtained from MD simulations with the TIP4P-Ew model, they are centered at larger $\langle R_{Cs^+-O_i} \rangle$ values, with their maxima shifted by $\sim$0.5 Å relative to the TIP4P-Ew distribution, and more uniform, with the difference in σ_i^2 between the narrowest and broadest *i*-RDFs being $\sim$0.5 Å^2 compared to $\sim$0.75 Å^2 for the corresponding TIP4P-Ew *i*-RDFs.

Finally, the ability of the different models to realistically describe the hydration structure of Cs^+ in solution is determined from comparisons between calculated and measured L_1-edge and L_3-edge EXAFS spectra shown in Fig. 4a-e and Fig. 4f-j, respectively. Two different methods are used to calculate the EXAFS spectra from the atomic positions of molecular configurations

extracted from either MD or PIMD trajectories. The first method (solid line) uses more accurate representations of the Cs^+ and O muffin tin potentials within the FEFF9 EXAFS calculation by selecting only the Cs^+ and O positions (i.e., H atoms are not included in the calculations), while the second method (dashed lines) employs the default settings that also includes the H positions (see Supporting Information for specific details).[247] The TIP4P-Ew model provides the poorest performance in reproducing both L_1-edge and L_3-edge spectra, predicting more pronounced oscillations whose amplitudes are too high and that are slightly out of phase with the experiment at mid- to high-*k* values. This reflects the narrow width of the first peak in the Cs^+–O RDF and the shorter Cs^+–O distance predicted by the TIP4P-Ew model (Fig. 2a). The inclusion of an implicit representation of many-body effects through classical polarization as implemented in the TTM-nrg model results in EXAFS spectra that more closely follow the oscillations in the corresponding experimental traces, although the calculated amplitudes of the oscillations are somewhat lower than in the measurements. The low amplitude and the frequency mismatch for the TTM-nrg model are progressively improved as 2B and 3B effects in the underlying Cs^+–water interactions are taken explicitly into account in the simulations with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models, which lead to excellent agreement with the experimental data for both L_1-edge and L_3-edge spectra. As expected from the analysis of the RDFs shown in Fig. 2, nuclear quantum effects appear to have only minimal impact on the EXAFS spectra.

In conclusion, the role played by many-body effects in the hydration structure of Cs^+ has been investigated through a systematic analysis of predictions obtained from MD and PIMD simulations carried out with four different molecular models built upon a hierarchy of approximate representations of Cs^+–water interactions. As expected from the simpler functional form, the empirical, pairwise-additive TIP4P-Ew model provides the poorest agreement with the available EXAFS spectra, which suggests a physically incorrect description of the underlying hydration structure. Although inclusion of many-body effects through an implicit representation based on classical polarization as implemented in the TTM-nrg model improves the agreement

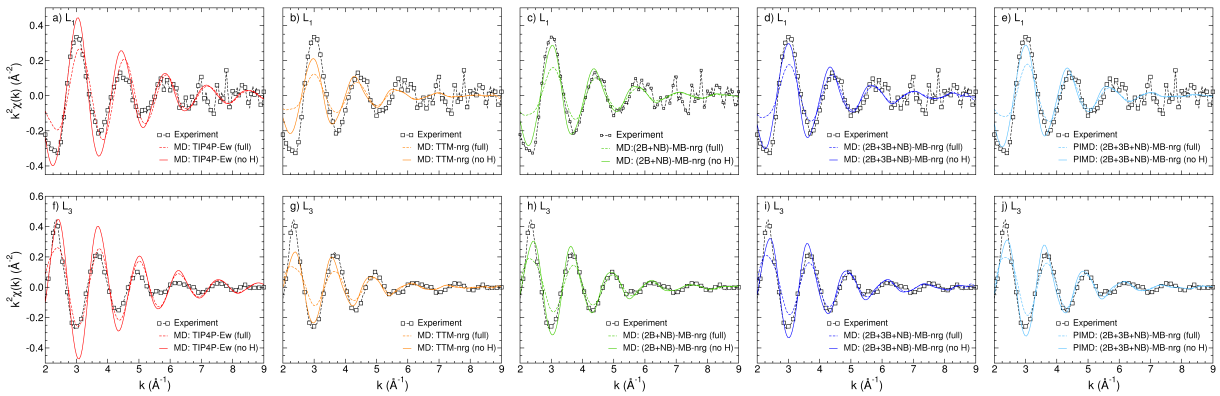

Figure 4.4. Comparisons between L_1-edge (panels a-e) and L_3-edge (panels f-j) EXAFS spectra, $k^2\chi(k)$, calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg model with the corresponding experimental data. See main text for details. Two sets of theoretical spectra are shown, corresponding to FEFF9 calculations carried out on molecular configurations that include only Cs^+ and O atoms (solid lines) and all atoms (dashed lines). See main text for details.

with experiment, the predicted Cs^+ hydration structure appears to lack the correct sequence of well-defined water shells. Systematic inclusion of explicit many-body Cs^+–water interactions in the MB-nrg models as well as account of nuclear quantum effects in PIMD simulations progressively leads to nearly quantitative agreement with the experimental EXAFS spectra. These results thus emphasize the importance of a physically correct representation, at both short and long ranges, of many-body effects in aqueous solutions, as suggested from early measurements of vibration-rotation spectra of water clusters.[242–245] Furthermore, the present analysis also indicates that PEFs rigorously built upon many-body expansions of interaction energies hold great promise for "realistic" simulations of aqueous solutions.

4.1 Acknowledgements

This research was supported by the National Science Foundation through grant CHE-1453204 and used resources of the Extreme Science and Engineering Discovery Environment (XSEDE),[?] which is supported by the National Science Foundation through grant ACI-1053575, under allocation TG-CHE110009. M.R.R. was supported by a Software Fellowship from the

Molecular Sciences Software Institute, which is funded by the National Science Foundation through grant ACI-1547580. Work by J.L.F. and G.K.S. was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences & Biosciences. Pacific Northwest National Laboratory (PNNL) is a multiprogram national laboratory operated for DOE by Battelle. The Cs^+ L_3 XAFS measurements were performed at the PHOENIX beamline of the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. The Cs^+ L_1 measurements at Advanced Photon Source, Sector 20 were supported by the U.S. Department of Energy and the Canadian Light Source.

Chapter 4, in full, is a reprint of the material as it appears in "Many-Body Effects Determine the Local Hydration Structure of Cs^+ in Solution", Zhuang, D.; Riera, M.; Schenter, G.K.; Fulton, J.L; Paesani, F.;J. Phys. Chem. Lett., 10, 406, 2019. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Development of an automated protocol to obtain and use MB-nrg PEFs.

5.1 Introduction

In previous chapters it was shown that the MB-nrg PEFs are able to accurately predict experimental properties of systems from the gas to the condensed phase. The process of developing, implementing, and using these MB-nrg PEFs has taken several years and, consequently, assuming that the same amount of time will be required to develop a PEF for any other system is not ideal. For this reason, a part of the time dedicated to this thesis has been to develop a software infrastructure that allows for the automatic generation of a PEF for any system at the desired level of theory, and allow the user to easily implement it and use it to run any kind of molecular dynamics simulations.

The protocol for obtaining a new PEF before this software package, which will be called MB-fit, involved a combination of multiple scripts and codes in different languages, and required an advanced knowledge of the methodology and advanced coding skills. Once the PEFs were generated, they had to be implemented in our unpublished in-house modified version of DL_POLY2,[248] which was originally hardcoded for MB-pol, making it difficult to implement new and more complex PEFs. The rigidness of the source code of DL_POLY2 does not allow for meaningful improvement of the efficiency, as it is not modular and is written in FORTRAN77, a very restrictive language. Thus, to bypass these issues, a software package designed to efficiently

use these MB-nrg PEFs in modern C++ has been written from scratch.

5.2 Generation of the MB-nrg PEFs

In order to use the MB-nrg PEFs for a new system they must first be generated through an automated process. For this reason, with the help of undergraduate students Ethan Bull-Vulpe and Kaushik Ganapathy, a Python library that allows non-expert users to obtain and generate PEFs was written, and guides them through the process of implementing them into the software package that calculates the energy. This library is a public github repository, and while still under development, the main features and functionalities are properly working. It has been used to generate test PEFs for molecules such as ethanol, acetonitrile, and molecular ions such as NH_4^+ and PF_6^-. The following sections will review the different steps that the library follows in order to generate new PEFs. All the steps are clearly documented with examples in a Jupyter notebook.[249]

5.2.1 Generation of the training set configurations

Machine learning (ML) techniques are extremely powerful, with the drawback that they usually need thousands of points in the training set.[250] The proper selection of these points is crucial in ensuring the success of describing the full PES with ML techniques. For each data point we add to the training set, i.e. each configuration of a monomer, dimer, trimer, etc, the energy must be calculated at a high level of theory. This calculation can become prohibitively expensive if the system has too many atoms. Thus, the number of points that we calculate should be large enough to avoid any overfitting and be a good representation of the PES, but not so large that the process of obtaining the energies is too time consuming. The procedure implemented allows the user to easily obtain a good sampling over the relevant phase space, although a new methodology based on Gaussian processes[251] (GP) is being developed by Yaoguang Zhai, a graduate student in our lab co-advised by Sicun Gao. This GP implementation will filter the configurations generated by the notebook before the energy is calculated and decide if that

configuration will be meaningful or not, reducing the number of electronic structure calculations that need to be performed.

One-body configurations

One-body distortions are obtained by performing a uniform normal mode sampling on the normal mode displacements of a the molecular species of interest. This sampling method is detailed in Ref. 252. Assuming that the system can be described within the harmonic approximation, the Hamiltonian can be represented as

$$\widehat{H}_h(T) = -\frac{\hbar^2}{2}\nabla^T \mathbf{M}^{-1}\nabla + \frac{1}{2}(\mathbf{r}-\mathbf{q})^T \mathbf{K}(\mathbf{r}-\mathbf{q}) + V_0 \tag{5.1}$$

where $\mathbf{r}$ is the coordinate vector, $\mathbf{M} = \mathrm{diag}(m_i)$ with m_i being the mass of each atom, V_0 is the effective harmonic potential, $\mathbf{q}$ its center, and $\mathbf{K}$ the Hessian matrix. Thus, the ensemble average, represented by

$$\langle f(\hat{\mathbf{r}})\rangle_h := \mathrm{Tr}[e^{-\beta\widehat{H}_h} f(\hat{\mathbf{r}})] = \| 2\pi\mathbf{D} \|^{-1/2} \int d\mathbf{x} e^{-\frac{1}{2}\mathbf{x}^T\mathbf{D}^{-1}\mathbf{x}} f(\mathbf{q}+\mathbf{x}) \tag{5.2}$$

with the displacement-displacement correlation matrix $\mathbf{D}$ defined as

$$\mathbf{D} = \mathbf{M}^{-1/2} d(\Omega) \mathbf{M}^{-1/2} \tag{5.3}$$

and the frequency matrix Ω given by

$$\Omega^2 = \tilde{\mathbf{K}} = \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2} \tag{5.4}$$

is used to sample configurations at a given temperature T.

By default, the sampling temperature smoothly transitions from the excitation temperature of the lowest mode to the temperature of the largest mode. At low temperatures, the low energy modes will be properly sampled but the high energy modes will not be sampled, while at large

temperatures, motion along the high energy modes will become accessible.

Two-body configurations

Two body configurations are more complex to sample than one-body, due to the higher number of degrees of freedom of the system. However, we can arbitrarily reduce it to 7 intermolecular degrees of freedom: the distance between the center of mass of both monomers, and two rigid rotations (three degrees of freedom each), one per monomer. These coordinates allow for good sampling of all possible configurations between the two monomers, ignoring the internal degrees of freedom.

The two-body training set configurations are obtained as a combination of three different sets of configurations. As in the one-body training set, the first is the normal-mode uniform sampling of all the minimum energy structures and transition states of the dimer, from now on, S_{NM}. In this case, the sampling is performed at low temperature to provide physically meaningful configurations. The second set of configurations are generated by rigid uniform scans with random rotations of the optimized monomers, S_r. The distance between the center of mass of the monomers is smoothly increased until a cutoff distance set by the user is reached. At each distance, a random rigid rotation is applied to each monomer, ensuring a proper sampling of all possible orientations at each distance if the number of configurations is large enough for that given system. Finally, since the polynomials also account for intramolecular distances, some configurations that contain distorted monomers at any distance between the center of mass also need to be added. To do so, a set of distorted monomers is generated for each monomer using the uniform normal-mode sampling procedure detailed above (S^1_{NM} and S^2_{NM}). The sampling must be done at low temperature, since at the usual simulation temperatures (0-300 K) monomer distortion energies higher than 40 kcal/mol are unlikely to be sampled. A similar procedure as for S_r is followed, but instead of using the minimum energy structure for the scan, a random distorted configuration is selected from S^1_{nm} and S^2_{NM}, and the same smooth increase of the distance between the center of mass of both monomers and corresponding rotations are performed. This

yields the flexible set of configurations, S_f. Thus, the total training set configurations, S_{ts} is the union of all the individual sets generated:

$$S_{ts} = S_{NM} \cup S_r \cup S_f \tag{5.5}$$

Three-body configurations

The previous chapters show that a good description of the three-body interactions is important to recover the proper behavior of the system and be able to predict and reproduce experimental results. However, an automated protocol to obtain the three-body training sets and polynomials has not yet been developed. Chapter 4 introduced the three-body PEF for $Cs^+(H_2O)_2$, which was obtained using the old methodology, i.e. using the collection of scripts written in various languages to generate the PEF and requiring familiarity with the different programs to implement it. Ideally, a smooth scan along the intermolecular degrees of freedom as detailed in the previous section should be performed, but since there is a high dimentionality for three or more monomers, this becomes unfeasible. Until the GP method is implemented and working, a temporary solution has been used to obtain a reasonable and physically meaningful training set. Once the two-body PEFs are obtained, tested, and used in simulations, configurations from the trajectories of these simulations are extracted and used as training set configurations for the three-body PEF.

5.2.2 Energy calculation and generation of the training set

Once the configurations are generated, the n-body energy of each one must be calculated. This step can be complicated if the user is not familiar with any quantum chemistry packages. The library uses a database structure, SQLite,[253] which manages the energy calculation automatically. The current implementation of MB-fit supports two different quantum chemistry packages, QChem[178] and Psi4,[254,255] with the user deciding which to use. The database handler will add the configuration to the database, decompose the system into its fragments, get the energy of

each fragment, and from there reconstruct the n-body energies. However, SQLite is not an efficient database structure for data containing more than the thousand entries and becomes highly inefficient when retrieving or adding information. For this reason, current work in progress includes migrating from SQLite to PostgreSQL.[256] This new database will provide a number of improvements, including more efficient reading and writing to the database, and more importantly, will enable distributed cloud computing, since it allows simultaneous writing into it.

5.2.3 Fitting

Once the training set generation is completed, and the energies of each configuration have been calculated, the coefficients of the polynomials and the nonlinear parameters in the variables need to be fitted so that the energy predicted is as close as possible to the reference energy. The first step for the fit is to generate the permutationally invariant polynomials that correspond to the system. MB-fit contains functions to generate the polynomials from a given monomer or dimer. This generation is preformed as follows.

1. The symmetry of the monomer needs to be defined. This symmetry is not the point group symmetry but the chemical equivalence of the atoms in the monomer. As an example, methane, CH_4, would have the symmetry tag A1B4, since the four hydrogen atoms are equivalent, thus permuting any of those atoms should give the exact same energy. In the same way, a dimer such as CO_2-CO_2 would have a symmetry tag A1B2-A1B2. With this symmetry known, the code generates all possible distances between all atoms in the system, which are assigned to variables. The program differentiates between intramolecular and intermolecular variables in order to define the specifications of the polynomials that we want to generate. This is outputted to a file that is then inputted to the polynomial generation, and can be modified by adding filters to remove terms that fulfill certain criteria, such as maximum polynomial degree for a given variable. A degree k polynomial will have a maximum of k variables being multiplied in the same term.

2. After deciding a maximum degree of the polynomial (the number of terms scales exponentially with the degree), the evaluation of the polynomials is optimized by MAPLE.[257] This optimization minimizes the number of floating point operations, reducing the evaluation time, but yields extremely long files that will not be possible to vectorize automatically. The trade-off between optimizing and not vectorizing and vectorizing but not optimizing has not been studied yet, and it is object of current investigation. After the corresponding optimized templates are generated, a small C++ library that contains the polynomial evaluation is created.

3. The Python C++ code writer uses all previous information in a fitting routine. This fitting routine follows the same procedure as described in Chapters 1 and 2. Several fits are run and the best fit is kept as the final fit. The test set is then evaluated using the final parameters, and for now, the user must decide if that test evaluation is satisfactory.

These steps are separated in different calls to multiple functions that turns what was originally a tedious and long process into a simple one that can be performed by non-experts in the field. After obtaining the final parameters of a fit that properly reproduces the test set energies, a last function call generates all the pieces of code that need to be added to the energy calculator, MBX (for many-body expansion), and will enable any sort of calculation or MD simulation requiring energies and gradients.

5.3 MBX: an MB-nrg energy calculator

This section will describe the structure and main features of MBX, a C++ library that efficiently uses the MB-nrg PEFs. As mentioned earlier, the old software packages that used MB-nrg PEFs were lacking either efficiency or generalization, and their rigid structures made it difficult to try to improve them. Thus, this new software package was written having two main goals in mind: generalization and efficiency. Due to its modular structure, MBX is able to easily accept new PEFs with minor modification, and in the near future, one will have to only add files

generated by MB-fit to the code and recompile to use these new PEFs.

There are two sources of speedup in MBX. The first one was already available in DL_POLY2, which is the parallel implementation of the expensive routines. In computer science there are two main ways of parallelazing software. The first one is a single node parallelization (multithreading), which is the one that for now is implemented in MBX using OpenMP.[258] The second one, which is more complicated to implement in an efficient way, is a multinode parallelization. This kind of parallel scheme will be required when simulations with tens of thousands of molecules are run, but for now, we are focusing on smaller systems made of hundreds of molecules, for which single node parallelization scales well in MBX.

The second source of speedup is the vectorization of most of the calculations. Vectorized operations take advantage of the width of the register. Recent computers have a width of 256 bits, and the newest processors have a 512 bit register width. A double type variable needs 64 bit to be processed. This means that we can process four doubles in a 256 bit register, gaining a four times speedup. The only shortcoming is that, as it happens with GPU computing, the data must be stored in continuous blocks in memory, implying that there cannot be jumps in memory from one operation to the next, which, depending on the nature of the data used, are sometimes difficult to allocate.

MBX has a central `System` class, which is the gateway to input and output data into and from the system. This class can be cast from an external routine, set up with the monomer information and the coordinates, and be used to obtain energy and gradients for a given configuration. Once the coordinates are set into the system, it starts the calculation of the different energy contributions one by one: one-body, two-body, three-body, dispersion, and electrostatics. Each one of these parts have been optimized to perform quickly, but there is still some room for improvement in performance and efficiency such as reducing the spin time in parallel computations, or enabling the vectorization of the polynomial functions along with the previously performed optimization. Once the energy and gradients are calculated, they can be easily retrieved and used by the external code. Then they can be sent to an MD driver such as i-PI[259] by taking advantage

of the driver that has also been developed along with MBX.

Although MBX is still being rigorously tested, energy calculations for large gas phase systems are working properly, and the analytical gradients are correct for the systems implemented and some test PEFs obtained using the potential fitting library. However, periodic boundary conditions (PBC) are still under development. Future work will include implementing the PBC calculations, enabling calculation of the virial tensor to allow MD simulations in the isobaric-isothermal ensemble (NPT), and writing extended documentation about the structure of the software and a user manual with extensive examples.

Chapter 6

Conclusions

Water is ubiquitous in nature, and ions and molecules dissolved in it can have a key role in different important processes, either natural or industrial. A correct description of the PES is needed to characterize structural, thermodynamic and dynamical properties, as well as the behavior of electrolyte solutions and other liquids in different environments. Building upon MB-pol, which correctly describes water properties from the gas to the condensed phase, the TTM-nrg PEFs were developed for alkali metal ions and water, which have a quasi-perfect description of the water interactions, and the ion-water interactions are described with a classical Buckingham function. The TTM-nrg PEFs are fully derived from accurate *ab initio* data at the CCSD(T)/CBS level of theory. The accuracy of the TTM-nrg PEFs has been assessed through an extensive analysis of energies and structures of $M^+(H_2O)_n$ dimers, with M = Li, Na, K, Rb, and Cs, and n = 1-4. Comparisons with *ab initio* data suggest that these PEFs, although not perfect, provide an efficient representation of alkali metal ion - water interactions that correctly include many-body effects.

As an improvement to the TTM-nrg PEFs, an explicit correction to the two-body term (V^{2B}) was developed for the MB-nrg PEFs for $M^+(H_2O)$ dimers, with M = Li, Na, K, Rb, and Cs. The repulsion term that was described by the Born-Mayer function is replaced by a set of permutationally invariant polynomials that are fitted to exactly reproduce the reference coupled-cluster data, and smoothly switch to zero at distances larger than 8 Å. Since the many-body

contributions are calculated in the same way as in TTM-nrg and MB-pol, the MB-nrg PEFs are fully compatible with these other functions. The accuracy of MB-nrg has been assessed thorugh an extensive analysis of both interaction energies and harmonic frequencies for the five alkali metal ions. Comparisons between MB-nrg and other *ab initio* methods such as MP2 or DFT, as well as with classical polarizable force fields (AMOEBA and TTM-nrg) demonstrate that MB-nrg always provides, in all cases, the highest accuracy, quantitatively reproducing the corresponding coupled-cluster reference values.

The previously developed MB-nrg PEFs with corrections up to the two-body term (MB-nrg 2B+NB) have been used to further investigate the microhydration of alkali metal ions. A detailed analysis was presented, at both classical and quantum levels, of the temperature dependence of isomeric equilibria and vibrational spectra of $M^+(H_2O)_n$ clusters, with M = Li, Na, K, Rb, and Cs, and n = 1-3. This analysis has been performed by processing the trajectories from classical replica-exchange molecular dynamics (REMD), and quantum replica-exchange path-integral molecular dynamics (REPIMD). It has been shown that nuclear quantum effects play a negligible role in determining the relative stability of different isomers of clusters containing the lightest alkali metal ions, Li^+ and Na^+, but being increasingly influential as the ion size increases. The calculated anharmonic infrared spectra shows excellent agreement with published experimental data, showing that there exists the possibility to directly connect experimental measurements with computer simulations at the molecular level, which provides hope for the development of a consistent picture of ion hydration from the gas to the condensed phase.

As a further improvement to the MB-nrg 2B+NB, the three-body correction term has been added to the Cs^+-water interaction, giving rise to the MB-nrg 2B+3B+NB model. The role played by many-body effects in the hydration structure of Cs^+ was then investigated through a systematic analysis of predictions obtained from MD and PIMD simulations carried out with different molecular models, built upon a hierarchy of different approximate representations of Cs^+-water interactions. As expected, the higher the complexity of the model, the better the agreement between the experimental and the predicted EXAFS spectra. These results thus

emphasize the importance of a physically correct representation, at both short and long ranges, of many-body effects in aqueous solutions, as suggested from early measurements of vibration-rotation spectra of water clusters.[242–245] Furthermore, the present analysis also indicates that PEFs rigorously built upon many-body expansions of interaction energies hold great promise for "realistic" simulations of aqueous solutions. Future work will include studying surface propensity of different ions and molecules, calculating potential of mean forces (PMFs) for the pathway of the molecules from the vacuum to the bulk, and calculating sum-frequency generation (SFG) infrared spectra for these systems, in a similar way that has been done for water-air interfaces.[174,260]

After showing that the MB-nrg PEFs provide physical representations of the systems, we have developed a software infrastructure to automatically generate, implement, and use these MB-nrg PEFs. Although it is still under development, most of the features are functioning and tested. The user can decide which molecule and at which level of theory the PEF must be calculated, and through an automated process that generates the training set, computes the energies, and fits the potential, the user obtains a PEF for the target system, which can then be implemented in MBX, an efficient software package that calculates MB-nrg energies and gradients, which is interfaced with the MD driver i-Pi. This package enables the generation and application of MB-nrg PEFs for any complex small molecule, enabling the area of predictive representations in molecular dynamics simulations.

Bibliography

[1] Felix Franks. *Water: a matrix of life*. Royal Society of Chemistry, 2007.

[2] Anders Nilsson and Lars GM Pettersson. The structural origin of anomalous properties of liquid water. *Nat. Commun.*, 6:8998, 2015.

[3] Makoto Endo. Calcium ion as a second messenger with special reference to excitation-contraction coupling. *J. Pharm. Sci.*, 100(5):519–524, 2006.

[4] S.N. Orlov and P. Hamet. Intracellular monovalent ions as second messengers. *J. Memb. Biol.*, 210(3):161–172, 2006.

[5] R. Giles Harrison and Hannes Tammet. Ions in the terrestrial atmosphere and other solar system atmospheres. *Space Sci. Rev.*, 137(1-4):107–118, 2008.

[6] Katrianne Lehtipalo, Linda Rondo, Jenni Kontkanen, Siegfried Schobesberger, Tuija Jokinen, Nina Sarnela, Andreas Kúrten, Sebastian Ehrhart, Alessandro Franchin, Tuomo Nieminen, Francesco Riccobono, Mikko Sipilä, Taina Yli-Juuti, Jonathan Duplissy, Alexey Adamov, Lars Ahlm, Joao Almeida, Antonio Amorim, Federico Bianchi, Martin Breitenlechner, Josef Dommen, Andrew J. Downard, Eimear M. Dunne, Richard C. Flagan, Roberto Guida, Jani Hakala, Armin Hansel, Werner Jud, Juha Kangasluoma, Veli-Matti Kerminen, Helmi Keskinen, Jaeseok Kim, Jasper Kirkby, Agnieszka Kupc, Oona Kupiainen-Maatta, Ari Laaksonen, Michael J. Lawler, Markus Leiminger, Serge Mathot, Tinja Olenius, Ismael K. Ortega, Antti Onnela, Tuukka Petaja, Arnaud Praplan, Matti P. Rissanen, Taina Ruuskanen, Filipe D. Santos, Simon Schallhart, Ralf Schnitzhofer, Mario Simon, James N. Smith, Jasmin Trostl, Georgios Tsagkogeorgas, Antonio Tome, Petri Vaattovaara, Hanna Vehkamaki, Aron E. Vrtala, Paul E. Wagner, Christina Williamson, Daniela Wimmer, Paul M. Winkler, Annele Virtanen, Neil M. Donahue, Kenneth S. Carslaw, Urs Baltensperger, Ilona Riipinen, Joachim Curtius, Douglas R. Worsnop, and Markku Kulmala. The effect of acid–base clustering and ions on the growth of atmospheric nano-particles. *Nat. Commun.*, 7:11594, 2016.

[7] Martin Winter and Ralph J. Brodd. What are batteries, fuel cells, and supercapacitors? *Chem. Rev.*, 104(10):4245–4270, 2004.

[8] Isadore Zipkin. The inorganic composition of bones and teeth. In *Biological calcification: cellular and molecular aspects*, pages 69–103. Springer, 1970.

[9] Masakazu Nakao and David C Gadsby. Voltage dependence of na translocation by the na/k pump. *Nature*, 323(6089):628, 1986.

[10] EM Knipping, MJ Lakin, KL Foster, P Jungwirth, DJ Tobias, RB Gerber, D Dabdub, and BJ Finlayson-Pitts. Experiments and simulations of ion-enhanced interfacial chemistry on aqueous nacl aerosols. *Science*, 288(5464):301–306, 2000.

[11] Eladio M Knipping and Donald Dabdub. Impact of chlorine emissions from sea-salt aerosol on coastal urban ozone. *Environ. Sci. Technol.*, 37(2):275–284, 2003.

[12] WR Simpson, R von Glasow, K Riedel, P Anderson, P Ariya, J Bottenheim, J Burrows, LJ Carpenter, U Frieß, Michael Evan Goodsite, et al. Halogens and their role in polar boundary-layer ozone depletion. *Atmospheric Chem. Phys.*, 7(16):4375–4418, 2007.

[13] Jack G Calvert, Allan Lazrus, Gregory L Kok, Brian G Heikes, James G Walega, John Lind, and Christopher A Cantrell. Chemical mechanisms of acid generation in the troposphere. *Nature*, 317(6032):27, 1985.

[14] Y. Marèchal. *The Hydrogen Bond and the Water Molecule: The Physics and Chemistry of Water, Aqueous and Bio-Media*. Elsevier, Amsterdam, 2006.

[15] Philip Ball. Water as an active constituent in cell biology. *Chem. Rev.*, 108:74–108, 2008.

[16] Pierandrea Lo Nostro and Barry W. Ninham. Hofmeister phenomena: An update on ion specificity in biology. *Chem. Rev.*, 112(4):2286–2322, 2012.

[17] Iván Restrepo-Angulo, Andrea De Vizcaya-Ruiz, and Javier Camacho. Ion channels in toxicology. *J. Appl. Toxicol.*, 30(6):497–512, 2010.

[18] Yizhak Marcus. Effect of ions on the structure of water: Structure making and breaking. *Chem. Rev.*, 109(3):1346–1370, 2009.

[19] Lukasz Piatkowski, Zhen Zhang, Ellen H. G. Backus, Huib J. Bakker, and Mischa Bonn. Extreme surface propensity of halide ions in water. *Nat. Commun.*, 5:4083, 2014.

[20] Valentina Migliorati, Francesco Sessa, Giuliana Aquilanti, and Paola D'Angelo. Unraveling halide hydration: A high dilution approach. *J. Chem. Phys.*, 141(4):044509–1–11, 2014.

[21] David C Gadsby, Junko Kimura, and Akinori Noma. Voltage dependence of na/k pump current in isolated heart cells. *Nature*, 315(6014):63, 1985.

[22] Douglas J. Tobias, Abraham C. Stern, Marcel D. Baer, Yan Levin, and Christopher J. Mundy. Simulation and theory of ions at atmospherically relevant aqueous liquid-air interfaces. *Annu. Rev. Phys. Chem.*, 64(1):339–359, 2013.

[23] David E. Smith and Liem X. Dang. Computer simulations of nacl association in polarizable water. *J. Chem. Phys.*, 100(5):3757–3766, 1994.

[24] Anne Willem Omta, Michel F. Kropman, Sander Woutersen, and Huib J. Bakker. Negligible effect of ions on the hydrogen-bond structure in liquid water. *Science*, 301(5631):347–349, 2003.

[25] Daniel Spångberg, Rossend Rey, James T. Hynes, and Kersti Hermansson. Rate and mechanisms for water exchange around li^+(aq) from md simulations. *J. Phys. Chem. B*, 107(18):4470–4477, 2003.

[26] Michel F. Kropman and Huib J. Bakker. Effect of ions on the vibrational relaxation of liquid water. *J. Am. Chem. Soc.*, 126(29):9135–9141, 2004.

[27] Jared D. Smith, Richard J. Saykally, and Phillip L. Geissler. The effects of dissolved halide anions on hydrogen bonding in liquid water. *J. Am. Chem. Soc.*, 129(45):13847–13856, 2007.

[28] S. Park and M.D Fayer. Hydrogen bond dynamics in aqueous nabr solutions. *Proc. Natl. Acad. Sci. USA*, 104:16731–16738, 2007.

[29] Michael D. Fayer, David E. Moilanen, Daryl Wong, Daniel E. Rosenfeld, Emily E. Fenn, and Sungnam Park. Water dynamics in salt solutions studied with ultrafast two-dimensional infrared (2d ir) vibrational echo spectroscopy. *Acc. Chem. Res.*, 42(9):1210–1219, 2009.

[30] David L. Bostick and Charles L. Brooks. Selective complexation of k^+ and na^+ in simple polarizable ion-ligating systems. *J. Am. Chem. Soc.*, 132(38):13185–13187, 2010.

[31] K. J. Tielrooij, N. Garcia-Araez, M. Bonn, and H. J. Bakker. Cooperativity in ion hydration. *Science*, 328(5981):1006–1009, 2010.

[32] Stefan Funkner, Gudrun Niehues, Diedrich A. Schmidt, Matthias Heyden, Gerhard Schwaab, Karen M. Callahan, Douglas J. Tobias, and Martina Havenith. Watching the low-frequency motions in aqueous salt solutions: The terahertz vibrational signatures of hydrated ions. *J. Am. Chem. Soc.*, 134(2):1030–1035, 2012.

[33] Phillip L. Geissler. Water interfaces, solvation, and spectroscopy. *Annu. Rev. Phys. Chem.*, 64(1):317–337, 2013.

[34] Pavel Jungwirth and Douglas J Tobias. Molecular structure of salt solutions: a new view of the interface with implications for heterogeneous atmospheric chemistry. *J. Phys. Chem. B*, 105(43):10468–10472, 2001.

[35] Pavel Jungwirth and Douglas J. Tobias. Ions at the air/water interface. *J. Phys. Chem. B*, 106(25):6361–6373, 2002.

[36] Liem X Dang and Tsun-Mei Chang. Molecular mechanism of ion binding to the liquid/vapor interface of water. *J. Phys. Chem. B*, 106(2):235–238, 2002.

[37] Liem X Dang. Computational study of ion binding to the liquid interface of water. *J. Chem. Phys. B*, 106(40):10388–10394, 2002.

[38] Poul B Petersen and Richard J Saykally. Confirmation of enhanced anion concentration at the liquid water surface. *Chem. Phys. Lett.*, 397(1-3):51–55, 2004.

[39] Poul B. Petersen and Richard J. Saykally. Probing the interfacial structure of aqueous electrolytes with femtosecond second harmonic generation spectroscopy. *J. Phys. Chem. B*, 110(29):14060–14073, 2006.

[40] Dingfang Liu, Gang Ma, Lori M Levering, and Heather C Allen. Vibrational spectroscopy of aqueous sodium halide solutions and air- liquid interfaces: Observation of increased interfacial depth. *The Journal of Physical Chemistry B*, 108(7):2252–2260, 2004.

[41] Elizabeth A Raymond and Geraldine L Richmond. Probing the molecular structure and bonding of the surface of aqueous salt solutions. *J. Phys. Chem. B*, 108(16):5051–5059, 2004.

[42] Bernd Winter, Ramona Weber, Philipp M Schmidt, Ingolf V Hertel, Manfred Faubel, Luboš Vrbka, and Pavel Jungwirth. Molecular structure of surface-active salt solutions: photoelectron spectroscopy and molecular dynamics simulations of aqueous tetrabutylammonium iodide. *J. Phys. Chem. B*, 108(38):14558–14564, 2004.

[43] Sutapa Ghosal, John C Hemminger, Hendrik Bluhm, Bongjin Simon Mun, Eleonore LD Hebenstreit, Guido Ketteler, D Frank Ogletree, Felix G Requejo, and Miquel Salmeron. Electron spectroscopy of aqueous solution interfaces reveals surface enhancement of halides. *Science*, 307(5709):563–566, 2005.

[44] Tsun-Mei Chang and Liem X Dang. Recent advances in molecular simulations of ion solvation at liquid interfaces. *Chem. Rev.*, 106(4):1305–1322, 2006.

[45] Poul B Petersen and Richard J Saykally. On the nature of ions at the liquid water surface. *Annu. Rev. Phys. Chem.*, 57:333–364, 2006.

[46] JA Pople, R Seeger, and R Krishnan. Variational configuration interaction methods and comparison with perturbation theory. *Int. J. Quantum Chem.*, 12(S11):149–163, 1977.

[47] JA Pople, R Krishnan, HB Schlegel, and JS Binkley. Electron correlation theories and their application to the study of simple reaction potential surfaces. *Int. J. Quantum Chem.*, 14(5):545–560, 1978.

[48] George D Purvis III and Rodney J Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76(4):1910–1918, 1982.

[49] Chr Møller and Milton S Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, 46(7):618, 1934.

[50] Martin Head-Gordon, John A Pople, and Michael J Frisch. Mp2 energy evaluation by direct methods. *Chem. Phys. Lett.*, 153(6):503–506, 1988.

[51] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864, 1964.

[52] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965.

[53] Robert G Parr. Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry*, pages 5–15. Springer, 1980.

[54] Siewert J Marrink, H Jelger Risselada, Serge Yefimov, D Peter Tieleman, and Alex H De Vries. The martini force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B*, 111(27):7812–7824, 2007.

[55] Valeria Molinero and Emily B Moore. Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B*, 113(13):4008–4016, 2008.

[56] Jibao Lu, Yuqing Qiu, Riccardo Baron, and Valeria Molinero. Coarse-graining of tip4p/2005, tip4p-ew, spc/e, and tip3p to monatomic anisotropic water models using relative entropy minimization. *J. Chem. Theory Comput.*, 10(9):4104–4120, 2014.

[57] Carlos Vega and Jose LF Abascal. Simulating water with rigid non-polarizable models: a general perspective. *Physical Chemistry Chemical Physics*, 13(44):19663–19688, 2011.

[58] Liem X Dang and B Montgomery Pettitt. Simple intramolecular model potentials for water. *Journal of Physical Chemistry*, 91(12):3349–3354, 1987.

[59] Pavel Jungwirth and Douglas J. Tobias. Specific ion effects at the air/water interface. *Chem. Rev.*, 106(4):1259–1281, 2006.

[60] Guillaume Lamoureux and Benoît Roux. Absolute hydration free energy scale for alkali and halide ions established from simulations with a polarizable force field. *J. Phys. Chem. B*, 110(7):3308–3322, 2006.

[61] Arieh Warshel, Mitsunori Kato, and Andrei V. Pisliakov. Polarizable force fields: History, test cases, and prospects. *J. Chem. Theory Comput.*, 3(6):2034–2045, 2007.

[62] Tiefeng Peng, Tsun-Mei Chang, Xiuquan Sun, Anh V. Nguyen, and Liem X. Dang. Development of ions-tip4p-ew force fields for molecular processes in bulk and at the aqueous interface using molecular simulations. *J. Mol. Liq.*, 173:47–54, 2012.

[63] Yue Shi, Zhen Xia, Jiajing Zhang, Robert Best, Chuanjie Wu, Jay W. Ponder, and Pengyu Ren. Polarizable atomic multipole-based amoeba force field for proteins. *J. Chem. Theory Comput.*, 9(9):4046–4063, 2013.

[64] Péter T. Kiss and András Baranyai. A new polarizable force field for alkali and halide ions. *J. Chem. Phys.*, 141(11):114501, 2014.

[65] Miriam Kohagen, Eva Pluhařová, Philip E. Mason, and Pavel Jungwirth. Exploring ion–ion interactions in aqueous solutions by a combination of molecular dynamics and neutron scattering. *J. Phys. Chem. Lett.*, 6:1563–1567, 2015.

[66] Steven J Stuart and BJ Berne. Effects of polarizability on the hydration of the chloride ion. *J. Phys. Chem.*, 100(29):11934–11943, 1996.

[67] G Lee Warren and Sandeep Patel. Comparison of the solvation structure of polarizable and nonpolarizable ions in bulk water and near the aqueous liquid- vapor interface. *J. Phys. Chem. C*, 112(19):7455–7467, 2008.

[68] Haibo Yu, Troy W. Whitfield, Edward Harder, Guillaume Lamoureux, Igor Vorobyov, Victor M. Anisimov, Alexander D. MacKerell, and Benoît Roux. Simulating monovalent and divalent ions in aqueous solution using a drude polarizable force field. *J. Chem. Theory Comput.*, 6(3):774–786, 2010.

[69] Lalith Perera and Max L Berkowitz. Many-body effects in molecular dynamics simulations of na+ (h2o) n and cl-(h2o) n clusters. *J. Chem. Phys.*, 95(3):1954–1963, 1991.

[70] Lalith Perera and Max L Berkowitz. Structure and dynamics of cl-(h2o) 20 clusters: The effect of the polarizability and the charge of the ion. *J. Chem. Phys.*, 96(11):8288–8294, 1992.

[71] James Caldwell, Liem X Dang, and Peter A Kollman. Implementation of nonadditive intermolecular potentials by use of molecular dynamics: development of a water-water potential and water-ion cluster interactions. *J. Am. Chem. Soc.*, 112(25):9144–9147, 1990.

[72] Liem X Dang and David E Smith. Molecular dynamics simulations of aqueous ionic clusters using polarizable water. *J. Chem. Phys.*, 99(9):6950–6956, 1993.

[73] Mark E. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.

[74] Jianshu Cao and Gregory A Voth. The formulation of quantum statistical mechanics based on the feynman path centroid density. i. equilibrium properties. *J. Chem. Phys.*, 100(7):5093–5105, 1994.

[75] Xiaolu Cheng and Ryan P Steele. Efficient anharmonic vibrational spectroscopy for large molecules using local-mode coordinates. *J. Chem. Phys.*, 141(10):104105, 2014.

[76] Xiaolu Cheng, Justin J Talbot, and Ryan P Steele. Tuning vibrational mode localization with frequency windowing. *J. Chem. Phys*, 145(12):124112, 2016.

[77] Yimin Wang and Joel M. Bowman. *Ab initio* potential and dipole moment surfaces for water. II. Local-monomer calculations of the infrared spectra of water clusters. *J. Chem. Phys.*, 134(15):154510, 2011.

[78] Christoph R Jacob and Markus Reiher. Localizing normal modes in large molecules. *J. Chem. Phys.*, 130(8):084106, 2009.

[79] H. Kistenmacher, H. Popkie, and E. Clementi. Study of the structure of molecular complexes. ii. energy surfaces for a water molecule in the field of a sodium or potassium cation. *J. Chem. Phys.*, 58(4):1689–1699, 1973.

[80] Robert Bukowski, Krzysztof Szalewicz, Gerrit C. Groenenboom, and Ad van der Avoird. Predictions of the Properties of Water from First Principles. *Science*, 315(5816):1249–1252, 2007.

[81] Robert Bukowski, Krzysztof Szalewicz, Gerrit C. Groenenboom, and Ad van der Avoird. Polarizable interaction potential for water from coupled cluster calculations. I. Analysis of dimer potential energy surface. *J. Chem. Phys.*, 128(9):094313, 2008.

[82] Robert Bukowski, Krzysztof Szalewicz, Gerrit C. Groenenboom, and Ad van der Avoird. Polarizable interaction potential for water from coupled cluster calculations. II. Applications to dimer spectra, virial coefficients, and simulations of liquid water. *J. Chem. Phys.*, 128(9):094314, 2008.

[83] Yimin Wang, Benjamin C. Shepler, Bastiaan J. Braams, and Joel M. Bowman. Full-dimensional, *ab initio* potential energy and dipole moment surfaces for water. *J. Chem. Phys.*, 131(5):054511, 2009.

[84] Yimin Wang, Xinchuan Huang, Benjamin C. Shepler, Bastiaan J. Braams, and Joel M. Bowman. Flexible, *ab initio* potential, and dipole moment surfaces for water. I. Tests and applications for clusters up to the 22-mer. *J. Chem. Phys.*, 134(9):094509, 2011.

[85] Volodymyr Babin, Gregory R. Medders, and Francesco Paesani. Toward a Universal Water Model: First Principles Simulations from the Dimer to the Liquid Phase. *J. Phys. Chem. Lett.*, 3(24):3765–3769, 2012.

[86] Gregory R. Medders, Volodymyr Babin, and Francesco Paesani. A Critical Assessment of Two-Body and Three-Body Interactions in Water. *J. Chem. Theory Comput.*, 9(2):1103–1114, feb 2013.

[87] Volodymyr Babin, Claude Leforestier, and Francesco Paesani. Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, vrt spectrum, and second virial coefficient. *J. Chem. Theory Comput.*, 9(12):5395–5403, 2013.

[88] Volodymyr Babin, Gregory R. Medders, and Francesco Paesani. Development of a "first principles" water potential with flexible monomers. ii: Trimer potential energy surface, third virial coefficient, and small clusters. *J. Chem. Theory Comput.*, 10(4):1599–1607, 2014.

[89] Gregory R. Medders, Volodymyr Babin, and Francesco Paesani. Development of a "first-principles" water potential with flexible monomers. iii. liquid phase properties. *J. Chem. Theory Comput.*, 10(8):2906–2910, 2014.

[90] Gerardo Andrés Cisneros, Kjartan Thor Wikfeldt, Lars Ojamäe, Jibao Lu, Yao Xu, Hedieh Torabifard, Albert P. Bartók, Gábor Csányi, Valeria Molinero, and Francesco Paesani. Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.*, 116:7501, 2016. PMID: 27186804.

[91] D Hankins, JW Moskowitz, and FH Stillinger. Water molecule interactions. *J. Chem. Phys.*, 53(12):4544–4554, 1970.

[92] Ryan M Richard and John M Herbert. A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory. *J. Chem. Phys.*, 137(6):064113, 2012.

[93] Ryan M Richard, Ka Un Lao, and John M Herbert. Understanding the many-body expansion for large systems. i. precision considerations. *J. Chem. Phys.*, 141(1):014108, 2014.

[94] Ka Un Lao, Kuan-Yu Liu, Ryan M Richard, and John M Herbert. Understanding the many-body expansion for large systems. ii. accuracy considerations. *J. Chem. Phys.*,

144(16):164105, 2016.

[95] Sotiris S Xantheas. Ab initio studies of cyclic water clusters (h2o) n, n= 1–6. ii. analysis of many-body interactions. *J. Chem. Phys.*, 100(10):7523–7534, 1994.

[96] Sotiris S Xantheas. Cooperativity and hydrogen bonding network in water clusters. *Chem. Phys.*, 258(2-3):225–231, 2000.

[97] Albert Defusco, Daniel P Schofield, and Kenneth D Jordan. Comparison of models with distributed polarizable sites for describing water clusters. *Mol. Phys.*, 105(19-22):2681–2696, 2007.

[98] Matthew J Elrod and Richard J Saykally. Many-body effects in intermolecular forces. *Chem. Rev.*, 94(7):1975–1997, 1994.

[99] Sandeep K Reddy, Shelby C Straight, Pushp Bajaj, C Huy Pham, Marc Riera, Daniel R Moberg, Miguel A Morales, Chris Knight, Andreas W Götz, and Francesco Paesani. On the accuracy of the mb-pol many-body potential for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice. *J. Chem. Phys.*, 145(19):194504, 2016.

[100] Francesco Paesani. Getting the right answers for the right reasons: toward predictive molecular simulations of water with many-body potential energy functions. *Acc. Chem. Res.*, 49(9):1844–1851, 2016.

[101] Harry Partridge and David W. Schwenke. The determination of an accurate isotope dependent potential energy surface for water from extensive ab initio calculations and experimental data. *J. Chem. Phys.*, 106(11):4618–4639, 1997.

[102] C. J. Burnham, D. J. Anick, P. K. Mankoo, and G. F. Reiter. The vibrational proton potential in bulk liquid water and ice. *J. Chem. Phys.*, 128(15):154519–1–20, 2008.

[103] Jonàs Sala, Elvira Guàrdia, and Marco Masia. The polarizable point dipoles method with electrostatic damping: Implementation on a model system. *The Journal of chemical physics*, 133(23):234101, 2010.

[104] Bastiaan J Braams and Joel M Bowman. Permutationally invariant potential energy surfaces in high dimensionality. *International Reviews in Physical Chemistry*, 28(4):577–606, 2009.

[105] Daniel J. Arismendi-Arrieta, Marc Riera, Pushp Bajaj, Rita Prosmiti, and Francesco Paesani. i-ttm model for ab initio-based ion–water interaction potentials. 1. halide–water potential energy functions. *J. Phys. Chem. B*, 120(8):1822–1832, 2016. PMID: 26560189.

[106] Pushp Bajaj, Andreas W. Götz, and Francesco Paesani. Toward chemical accuracy in the description of ion-water interactions through many-body representations. i. halide-water dimer potential energy surfaces. *J. Chem. Theory Comput.*, 12(6):2698–2705, 2016. PMID: 27145081.

[107] Marc Riera, Andreas W Götz, and Francesco Paesani. The i-ttm model for ab initio-based ion–water interaction potentials. ii. alkali metal ion–water potential energy functions. *Phys. Chem. Chem. Phys.*, 18(44):30334–30343, 2016.

[108] Hans W. Horn, William C. Swope, Jed W. Pitera, Jeffry D. Madura, Thomas J. Dick, Greg L. Hura, and Teresa Head-Gordon. Development of an improved four-site water model for biomolecular simulations: Tip4p-ew. *J. Chem. Phys.*, 120(20):9665–9678, 2004.

[109] In Suk Joung and Thomas E Cheatham III. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B*, 113(40):13279–13290, 2009.

[110] Haegyeom Kim, Jihyun Hong, Kyu-Young Park, Hyungsub Kim, Sung-Wook Kim, and Kisuk Kang. Aqueous rechargeable li and na ion batteries. *Chem. Rev.*, 114(23):11788–11827, 2014.

[111] Ali A. Hassanali, Jérôme Cuny, Vincenzo Verdolino, and Michele Parrinello. Aqueous solutions: State of the art in ab initio molecular dynamics. *Phil. Trans. R. Soc. A*, 372(2011):20120482, 2014.

[112] Collin D. Wick, Liem X. Dang, and Pavel Jungwirth. Simulated surface potentials at the vapor-water interface for the kcl aqueous electrolyte solution. *J. Chem. Phys.*, 125(2):024706, 2006.

[113] Collin D. Wick, I-Feng W. Kuo, Christopher J. Mundy, and Liem X. Dang. The effect of polarizability for understanding the molecular structure of aqueous interfaces. *J. Chem. Theory Comput.*, 3(6):2002–2010, 2007.

[114] Aleksandr V. Marenich, Ryan M. Olson, Adam C. Chamberlin, Christopher J. Cramer, and Donald G. Truhlar. Polarization effects in aqueous and nonaqueous solutions. *J. Chem. Theory Comput.*, 3(6):2055–2067, 2007.

[115] Collin D. Wick and Sotiris S. Xantheas. Computational investigation of the first solvation shell structure of interfacial and bulk aqueous chloride and iodide ions. *J. Phys. Chem. B*, 113(13):4141–4146, 2009.

[116] Collin D. Wick. Electrostatic dampening dampens the anion propensity for the air-water interface. *J. Chem. Phys.*, 131(8):084715, 2009.

[117] Elvira Guàrdia, Ioannis Skarmoutsos, and Marco Masia. On ion and molecular polarization of halides in water. *J. Chem. Theory Comput.*, 5(6):1449–1453, 2009.

[118] Carl Caleman, Jochen S. Hub, Paul J. van Maaren, and David van der Spoel. Atomistic simulation of ion solvation in water explains surface preference of halides. *Proc. Natl. Acad. Sci. USA*, 108(17):6838–6842, 2011.

[119] Lu Sun, Xin Li, Yaoquan Tu, and Hans Agren. Origin of ion selectivity at the air/water interface. *Phys. Chem. Chem. Phys.*, 17:4311–4318, 2015.

[120] James M. Lisy. Spectroscopy and structure of solvated alkali-metal ions. *Int. Rev. Phys. Chem.*, 16(3):267–289, 1997.

[121] Jean-Joseph Max and Camille Chapados. Ir spectroscopy of aqueous alkali halide solutions: Pure salt-solvated water spectra and hydration numbers. *J. Chem. Phys.*, 115(6):2664–2675, 2001.

[122] Dorothy J. Miller and James M. Lisy. Hydrated alkali-metal cations: Infrared spectroscopy and ab initio calculations of m+(h2o)x=2-5ar cluster ions for m = li, na, k, and cs. *J. Am. Chem. Soc.*, 130(46):15381–15392, 2008. PMID: 18939843.

[123] Ismael A. Heisler and Stephen R. Meech. Low-frequency modes of aqueous alkali halide solutions: Glimpsing the hydrogen bonding vibration. *Science*, 327(5967):857–860, 2010.

[124] Alan Grossfield, Pengyu Ren, and Jay W. Ponder. Ion solvation thermodynamics from simulation with a polarizable force field. *J. Am. Chem. Soc.*, 125(50):15671–15682, 2003. PMID: 14664617.

[125] Jean-Philip Piquemal, Lalith Perera, G. Andrés Cisneros, Pengyu Ren, Lee G. Pedersen, and Thomas A. Darden. Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: From energetics to structure. *J. Chem. Phys.*, 125(5):054511, 2006.

[126] Johnny C. Wu, Jean-Philip Piquemal, Robin Chaudret, Peter Reinhardt, and Pengyu Ren. Polarizable molecular dynamics simulation of zn(ii) in water using the amoeba force field. *J. Chem. Theory Comput.*, 6(7):2059–2070, 2010. PMID: 21116445.

[127] Jin Yu Xiang and Jay W. Ponder. A valence bond model for aqueous cu(ii) and zn(ii) ions in the amoeba polarizable force field. *J. Comp. Chem.*, 34(9):739–749, 2013.

[128] Aude Marjolin, Christophe Gourlaouen, Carine Clavaguéra, Pengyu Y. Ren, Johnny C. Wu, Nohad Gresh, Jean-Pierre Dognon, and Jean-Philip Piquemal. Toward accurate solvation dynamics of lanthanides and actinides in water using polarizable force fields: from gas-phase energetics to hydration free energies. *Theor. Chem. Acc.*, 131(4):1–14,

2012.

[129] Gregory R. Medders and Francesco Paesani. Infrared and raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics. *J. Chem. Theory Comput.*, 11(3):1145–1154, 2015.

[130] Thomas B. Adler, Gerald Knizia, and Hans Joachim Werner. A simple and efficient ccsd(t)-f12 approximation. *J. Chem. Phys.*, 127:221106, 2007.

[131] Gerald Knizia, Thomas B. Adler, and Hans Joachim Werner. Simplified ccsd(t)-f12 methods: Theory and benchmarks. *J. Chem. Phys.*, 130:054104, 2009.

[132] Urszula Góra, Rafał Podeszwa, Wojciech Cencek, and Krzysztof Szalewicz. Interaction energies of large clusters from many-body expansion. *J. Chem. Phys.*, 135(22):224102, 2011.

[133] J. Grant Hill, Kirk A. Peterson, Gerald Knizia, and Hans-Joachim Werner. Extrapolating mp2 and ccsd explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets. *J. Chem. Phys.*, 131(19):194105, 2009.

[134] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90(2):1007–1023, 1989.

[135] Rick A. Kendall, Thom H. Dunning, and Robert J. Harrison. Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions. *J. Chem. Phys.*, 96(9):6796–6806, 1992.

[136] David E. Woon and Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. v. core-valence basis sets for boron through neon. *J. Chem. Phys.*, 103(11):4572–4585, 1995.

[137] Kirk Peterson. private communication.

[138] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7:3297–3305, 2005.

[139] S.F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Mol. Phys.*, 19(4):553–566, 1970.

[140] H.-J. Werner, P.J. Knowles, G. Knizia, R. Manby, M. Schütz, and et al. Molpro, version 2012.1, a package of ab initio programs, 2012. see http://www.molpro.net.

[141] Hans-Joachim Werner, Peter J. Knowles, Gerald Knizia, Frederick R. Manby, and Martin Schütz. Molpro: A general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2(2):242–253, 2012.

[142] Eric M. Mas, Krzysztof Szalewicz, Robert Bukowski, and Bogumil Jeziorski. Pair potential for water from symmetry-adapted perturbation theory. *J. Chem. Phys.*, 107(11):4207–4218, 1997.

[143] David E. Woon and Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. iv. calculation of static electrical response properties. *J. Chem. Phys.*, 100(4):2975–2988, 1994.

[144] Ivan S. Lim, Jon K. Laerdahl, and Peter Schwerdtfeger. Fully relativistic coupled-cluster static dipole polarizabilities of the positively charged alkali ions from li+ to 119+. *J. Chem. Phys.*, 116(1):172–178, 2002.

[145] Yashpal Singh, B. K. Sahoo, and B. P. Das. Correlation trends in the ground-state static electric dipole polarizabilities of closed-shell atoms and ions. *Phys. Rev. A*, 88:062504, Dec 2013.

[146] Erin R. Johnson and Axel D. Becke. A post-hartree–fock model of intermolecular interactions. *J. Chem. Phys.*, 123(2):024101, 2005.

[147] Felix O. Kannemann and Axel D. Becke. van der waals interactions in density-functional theory: Intermolecular complexes. *J. Chem. Theory Comput.*, 6:1081–1088, 2010.

[148] A. Otero-de-la Roza and Erin R. Johnson. Non-covalent interactions and thermochemistry using xdm-corrected hybrid and range-separated hybrid density functionals. *J. Chem. Phys.*, 138:204109, 2013.

[149] Oleg A. Vydrov and G. E. Scuseria. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.*, 125:234109, 2006.

[150] Oleg A. Vydrov, J. Heyd, Aliaksandr V. Krukau, and G. E. Scuseria. Importance of short-range versus long-range hartree-fock exchange for the performance of hybrid density functionals. *J. Chem. Phys.*, 125:074106, 2006.

[151] Oleg A. Vydrov, G. E. Scuseria, and John P. Perdew. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.*, 126:154109, 2007.

[152] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai,

T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision d.01. Gaussian Inc. Wallingford CT 2009.

[153] Jay W. Ponder and Frederic M. Richards. An efficient newton-like method for molecular mechanics energy minimization of large molecules. *J. Comp. Chem.*, 8(7):1016–1024, 1987.

[154] Craig E. Kundrot, Jay W. Ponder, and Frederic M. Richards. Algorithms for calculating excluded volume and its derivatives as a function of molecular conformation and their use in energy minimization. *J. Comp. Chem.*, 12(3):402–409, 1991.

[155] Jay W. Ponder, Chuanjie Wu, Pengyu Ren, Vijay S. Pande, John D. Chodera, Michael J. Schnieders, Imran Haque, David L. Mobley, Daniel S. Lambrecht, Jr. Robert A. DiStasio, Martin Head-Gordon, Gary N. I. Clark, Margaret E. Johnson, and Teresa Head-Gordon. Current status of the amoeba polarizable force field. *The Journal of Physical Chemistry B*, 114(8):2549–2564, 2010. PMID: 20136072.

[156] Marie L. Laury, Lee-Ping Wang, Vijay S. Pande, Teresa Head-Gordon, and Jay W. Ponder. Revised parameters for the amoeba polarizable atomic multipole water model. *J. Phys. Chem. B*, 119(29):9423–9437, 2015. PMID: 25683601.

[157] Pengyu Ren and Jay W. Ponder. private communication.

[158] K. T. Tang and J. Peter Toennies. An improved simple model for the van der waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.*, 80(8):3726–3741, 1984.

[159] Jongseob Kim, Han Myoung Lee, Seung Bum Suh, D. Majumdar, and Kwang S. Kim. Comparative ab initio study of the structures, energetics and spectra of $x^-(h_2o)_{n=1-4}$ [x = f, cl, br, i] clusters. *J. Chem. Phys.*, 113(13):5259–5272, 2000.

[160] Jmol: an open-source java viewer for chemical structures in 3d. http://www.jmol.org/. Version 13.0.

[161] Gregory R. Medders, Andreas W. Götz, Miguel A. Morales, and Francesco Paesani. On the representation of many-body interactions in water. *J. Chem. Phys.*, 143(10):104102, 2015.

[162] Axel D. Becke. Density-functional exchange-enery approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098, 1988.

[163] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785, 1988.

[164] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 78:1396, 1996.

[165] J. M. Tao, John P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91:146401, 2003.

[166] Carlo Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *J. Chem. Phys.*, 110:6158, 1999.

[167] Axel D. Becke. Density functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98:5648, 1993.

[168] Jeng-Da Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128:084106, 2008.

[169] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.*, 10:6615, 2008.

[170] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *J. Chem. Phys.*, 132(15):154104, 2010.

[171] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, 32(7):1456–1465, 2011.

[172] Jeremy O. Richardson, Cristóbal Pérez, Simon Lobsiger, Adam A. Reid, Berhane Temelso, George C. Shields, Zbigniew Kisiel, David J. Wales, Brooks H. Pate, and Stuart C. Althorpe. Concerted hydrogen-bond breaking by quantum tunneling in the water hexamer prism. *Science*, 351(6279):1310–1313, 2016.

[173] William T. S. Cole, James D. Farrell, David J. Wales, and Richard J. Saykally. Structure and torsional dynamics of the water octamer from THz laser spectroscopy near 215 μm. *Science*, 352(6290):1194–1197, 2016.

[174] Gregory R. Medders and Francesco Paesani. Dissecting the molecular structure of the

air/water interface from quantum simulations of the sum-frequency generation spectrum. *J. Am. Chem. Soc.*, 138(11):3912–3919, 2016.

[175] C. Huy Pham, Sandeep K. Reddy, Karl Chen, Chris Knight, and Francesco Paesani. Many-body interactions in ice. *Journal of Chemical Theory and Computation*, 13(4):1778–1784, 2017.

[176] Ivan S Lim, Peter Schwerdtfeger, Bernhard Metz, and Hermann Stoll. All-electron and relativistic pseudopotential studies for the group 1 element polarizabilities from k to element 119. *J. Chem. Phys.*, 122(10):104103, 2005.

[177] Narbe Mardirossian and Martin Head-Gordon. Mapping the genome of meta-generalized gradient approximation density functionals: The search for b97m-v. *J. Chem. Phys.*, 142(7):074111, 2015.

[178] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kús, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr., H. Dop, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, P. A. Pieniazek, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, N. Sergueev, S. M. Sharada, S. Sharmaa, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, V. Vanovschi, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhou, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xua, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon. Advances in molecular quantum chemistry contained in the q-chem 4 program package. *Mol. Phys.*, 113:184–215, 2015.

[179] Pengyu Ren, , and Jay W. Ponder. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B*, 107(24):5933–5947, 2003.

[180] Andrey Tikhonov. Solution of incorrectly formulated problems and the regularization method. In *Soviet Math. Dokl.*, volume 5, pages 1035–1038, 1963.

[181] Yingkai Zhang and Weitao Yang. Comment on "generalized gradient approximation made simple". *Phys. Rev. Lett.*, 80:890–890, Jan 1998.

[182] Arindam Bankura, Anwesa Karmakar, Vincenzo Carnevale, Amalendu Chandra, and Michael L. Klein. Structure, dynamics, and spectral diffusion of water from first-principles molecular dynamics. *The Journal of Physical Chemistry C*, 118(50):29401–29411, 2014.

[183] L. B. Skinner, M. Galib, J. L. Fulton, C. J. Mundy, J. B. Parise, V.-T. Pham, G. K. Schenter, and C. J. Benmore. The structure of liquid water up to 360 mpa from x-ray diffraction measurements using a high q-range and from molecular simulation. *J. Chem. Phys*, 144(13):134504, 2016.

[184] M. Galib, M. D. Baer, L. B. Skinner, C. J. Mundy, T. Huthwelker, G. K. Schenter, C. J. Benmore, N. Govind, and J. L. Fulton. Revisiting the hydration structure of aqueous na+. *J. Chem. Phys.*, 146(8):084504, 2017.

[185] Ondrej Marsalek and Thomas E. Markland. Quantum dynamics and spectroscopy of ab initio liquid water: The interplay of nuclear and electronic quantum effects. *The Journal of Physical Chemistry Letters*, 8(7):1545–1551, 2017. PMID: 28296422.

[186] Richard A Sneen. Substitution at a saturated carbon atom. xvii. organic ion pairs as intermediates in nucleophilic substitution and elimination reactions. *Acc. Chem. Res.*, 6(2):46–53, 1973.

[187] M.J. Pregel, E.J. Dunn, R. Nagelkerke, G.R.J. Thatcher, and E. Buncel. Alkali–metal lon catalysis and inhibition in nucleophilic displacement reaction of phosphorus–sulfur–and carbon– based esters. *Chem. Soc. Rev.*, 24(6):449–455, 1995.

[188] S Nahar and HA Tajmir-Riahi. Do metal ions alter the protein secondary structure of a light-harvesting complex of thylakoid membranes? *J. Inorg. Biochem.*, 58(3):223–234, 1995.

[189] Sarah A Woodson. Metal ions and rna folding: A highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.*, 9(2):104–109, 2005.

[190] David E Draper. Rna folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.*, 95(12):5489–5495, 2008.

[191] Gary Siuzdak, Yoshitaka Ichikawa, Thomas J Caulfield, Benito Munoz, Chi Huey Wong, and KC Nicolaou. Evidence of calcium^{2+}-dependent carbohydrate association through ion spray mass spectrometry. *J. Am. Chem. Soc.*, 115(7):2877–2881, 1993.

[192] Zhi-Jie Tan and Shi-Jie Chen. Ion-mediated nucleic acid helix-helix interactions. *Biophys. J.*, 91(2):518–536, 2006.

[193] Norbert Sträter, William N Lipscomb, Thomas Klabunde, and Bernt Krebs. Two-metal ion catalysis in enzymatic acyl-and phosphoryl-transfer reactions. *Angew. Chem. Int. Ed.*, 35(18):2024–2055, 1996.

[194] Roland K.O. Sigel and Anna Marie Pyle. Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. *Chem. Rev.*, 107(1):97–113, 2007.

[195] Raven Hanna and Jennifer A Doudna. Metal ions in ribozyme folding and catalysis. *Curr. Opin. Chem. Biol.*, 4(2):166–170, 2000.

[196] Mary R. Stahley and Scott A. Strobel. Rna splicing: Group i intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr. Opin. Struct. Biol.*, 16(3):319–326, 2006.

[197] H.J. Bakker. Structural dynamics of aqueous salt solutions. *Chem. Rev.*, 108(4):1456–1473, 2008.

[198] Travis P. Pollard and Thomas L. Beck. Toward a quantitative theory of hofmeister phenomena: From quantum effects to thermodynamics. *Curr. Opin. Colloid Interface Sci.*, 23:110 – 118, 2016.

[199] Michael J. Gillan, Dario Alfè, and Angelos Michaelides. Perspective: How good is dft for water? *J. Chem. Phys.*, 144(13):130901, 2016.

[200] Volodymyr Babin, Claude Leforestier, and Francesco Paesani. Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, vrt spectrum, and second virial coefficient. *J. Chem. Theory Comput.*, 9(12):5395–5403, 2013.

[201] Sandra E. Brown, Andreas W. Götz, Xiaolu Cheng, Ryan P. Steele, Vladimir A. Mandelshtam, and Francesco Paesani. Monitoring water clusters "melt" through vibrational spectroscopy. *J. Am. Chem. Soc.*, 139(20):7082–7088, 2017.

[202] Shelby C. Straight and Francesco Paesani. Exploring electrostatic effects on the hydrogen bond network of liquid water through many-body molecular dynamics. *J. Phys. Chem. B*, 120(33):8539, 2016.

[203] Sandeep K. Reddy, Daniel R. Moberg, Shelby C. Straight, and Francesco Paesani. Temperature-dependent vibrational spectra and structure of liquid water from classical and quantum simulations with the mb-pol potential energy function. *J. Chem. Phys.*, 147, 2017.

[204] Daniel R. Moberg, Shelby C. Straight, Christopher Knight, and Francesco Paesani. Molecular origin of the vibrational structure of ice ih. *J. Phys. Chem. Lett.*, 8(12):2579–2583, 2017.

[205] Alex P Gaiduk, Tuan Anh Pham, Marco Govoni, Francesco Paesani, and Giulia Galli. Electron affinity of liquid water. *Nat. Commun.*, 9(1):247, 2018.

[206] Marc Riera, Narbe Mardirossian, Pushp Bajaj, Andreas W. Götz, and Francesco Paesani. Toward chemical accuracy in the description of ion–water interactions through many-body representations. alkali-water dimer potential energy surfaces. *J. Chem. Phys.*, 147(16):161715, 2017.

[207] Eugene Kamarchik, Yimin Wang, and Joel M Bowman. Quantum vibrational analysis and infrared spectra of microhydrated sodium ions using an ab initio potential. *J. Chem. Phys.*, 134(11):114311, 2011.

[208] Pushp Bajaj, Xiao-Gang Wang, Tucker Carrington Jr., and Francesco Paesani. Vibrational Spectra of Halide-Water Dimers: Insights on Ion Hydration from Full-Dimensional Quantum Calculations on Many-Body Potential Energy Surfaces. *J. Chem. Phys.*, 148(10):102321, 2018.

[209] AW Castleman and KHa Bowen. Clusters: Structure, energetics, and dynamics of intermediate states of matter. *J. Phys. Chem.*, 100(31):12911–12944, 1996.

[210] Michael A Duncan. Spectroscopy of metal ion complexes: Gas-phase models for solvation. *Annu. Rev. Phys. Chem.*, 48(1):69–93, 1997.

[211] Johan Mähler and Ingmar Persson. A study of the hydration of the alkali metal ions in aqueous solution. *Inorg. Chem.*, 51(1):425–438, 2011.

[212] Song Hi Lee and Jayendran C Rasaiah. Molecular dynamics simulation of ion mobility. 2. alkali metal and halide ions using the spc/e model for water at 25 c. *J. Phys. Chem.*, 100(4):1420–1425, 1996.

[213] Kasper P Jensen and William L Jorgensen. Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J. Chem. Theory Comput.*, 2(6):1499–1509, 2006.

[214] In Suk Joung and Thomas E. Cheatham III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*, 112(30):9020–9041, 2008.

[215] Sameer Varma and Susan B Rempe. Coordination numbers of alkali metal ions in aqueous solutions. *Biophys. Chem.*, 124(3):192–199, 2006.

[216] Timothy D Vaden, Corey J Weinheimer, and James M Lisy. Evaporatively Cooled $M^+(H_2O)$ Ar Cluster Ions: Infrared Spectroscopy and Internal Energy Simulations. *J. Chem. Phys.*, 121(7):3102–3107, 2004.

[217] Timothy D. Vaden, James M. Lisy, Prosser D. Carnegie, E. Dinesh Pillai, and Michael A. Duncan. Infrared spectroscopy of the $li^+(h_2o)$ar complex: The role of internal energy and its dependence on ion preparation. *Phys. Chem. Chem. Phys.*, 8:3078–3082, 2006.

[218] Amy L Nicely, Dorothy J Miller, and James M Lisy. Gas-phase vibrational spectroscopy and ab initio calculations of $rb^+(h_2o)_n$ and $rb^+(h_2o)_n$ar cluster ions. *J. Mol. Spectros.*, 257(2):157–163, 2009.

[219] Franziska Schulz and Bernd Hartke. Structural information on alkali cation microhydration clusters from infrared spectra. *Phys. Chem. Chem. Phys.*, 5(22):5021–5030, 2003.

[220] Sowmianarayanan Rajamani, Tuhin Ghosh, and Shekhar Garde. Size dependent ion hydration, its asymmetry, and convergence to macroscopic behavior. *J. Chem. Phys.*, 120(9):4457–4466, 2004.

[221] Han Myoung Lee, P Tarakeshwar, Jungwon Park, Maciej Roman Kołaski, Yeo Jin Yoon, Hai-Bo Yi, Woo Youn Kim, and Kwang S Kim. Insights into the structures, energetics, and vibrations of monovalent cation-(water) 1-6 clusters. *J. Phys. Chem. A*, 108(15):2949–2958, 2004.

[222] Franziska Schulz and Bernd Hartke. A new proposal for the reason of magic numbers in alkali cation microhydration clusters. *Theor. Chem. Acc.*, 114(4-5):357–379, 2005.

[223] Yimin Wang, Volodymyr Babin, Joel M. Bowman, and Francesco Paesani. The Water Hexamer: Cage, Prism, or Both. Full Dimensional Quantum Simulations Say Both. *J. Am. Chem. Soc.*, 134(27):11116–11119, 2012.

[224] Frank Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons, 2017.

[225] Takehiko Shimanouchi. *Tables of Molecular Vibrational Frequencies, Consolidated Volume I*. National Bureau of Standards, 1 edition, 1972.

[226] Phillip E Fraley and K Narahari Rao. High resolution infrared spectra of water vapor: $\nu 1$ and $\nu 3$ band of $h_2^{16}o$. *J. Mol. Spectros.*, 29(1-3):348–364, 1969.

[227] Haochen Ke, Christian van der Linde, and James M Lisy. Insights into gas-phase structural conformers of hydrated rubidium and cesium cations, $m^+(h_2o)_n$ ar (m= rb, cs; n= 3–5), using infrared photodissociation spectroscopy. *J. Phys. Chem. A*, 118(8):1363–1373, 2014.

[228] J-M Tarascon and Michel Armand. Issues and challenges facing rechargeable lithium batteries. In *Materials For Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group*, pages 171–179. World Scientific, 2011.

[229] Mogens Schou. Forty years of lithium treatment. *Arch. Gen. Psychiat.*, 54(1):9–13, 1997.

[230] Simon V Avery. Microbial interactions with caesium – implications for biotechnology. *J. Chem. Technol. Biotechnol.*, 62(1):3–16, 1995.

[231] Silvia Ramos, GW Neilson, AC Barnes, and P Buchanan. An anomalous x-ray diffraction study of the hydration structures of cs^{+} and i^{-} in concentrated solutions. *J. Chem. Phys.*, 123(21):214501, 2005.

[232] Viktória Mile, László Pusztai, Hector Dominguez, and Orest Pizio. Understanding the structure of aqueous cesium chloride solutions by combining diffraction experiments, molecular dynamics simulations, and reverse monte carlo modeling. *J. Phys. Chem. B*, 113(31):10760–10769, 2009.

[233] Viktória Mile, Orsolya Gereben, Shinji Kohara, and László Pusztai. On the structure of aqueous cesium bromide solutions: Diffraction experiments, molecular dynamics simulations and reverse monte carlo modeling. *J. Mol. Liq.*, 157(1):36–42, 2010.

[234] Song Hi Lee and Jayendran C Rasaiah. Molecular dynamics simulation of ionic mobility. i. alkali metal cations in water at 25 ^{o}c. *J. Chem. Phys.*, 101(8):6964–6974, 1994.

[235] Christian F Schwenk, Thomas S Hofer, and Bernd M Rode. "structure breaking" effect of hydrated cs+. *J. Phys. Chem. A*, 108(9):1509–1514, 2004.

[236] Stephan Deublein, Jadran Vrabec, and Hans Hasse. A set of molecular models for alkali and halide ions in aqueous solution. *J. Chem. Phys.*, 136(8):084501, 2012.

[237] Takashi Ikeda and Mauro Boero. Communication: Hydration structure and polarization of heavy alkali ions: A first principles molecular dynamics study of rb^{+} and cs^{+}. *J. Chem. Phys.*, 137:041101, 2012.

[238] Daniel Z Caralampio, José M Martínez, Rafael R Pappalardo, and Enrique Sánchez Marcos. The hydration structure of the heavy-alkalines rb^{+} and cs^{+} through molecular dynamics and x-ray absorption spectroscopy: Surface clusters and eccentricity. *Phys. Chem. Chem. Phys.*, 19(42):28993–29004, 2017.

[239] Henry S. Frank and Wen-Yang Wen. Ion-solvent interaction. structural aspects of ion-solvent interaction in aqueous solutions: a suggested picture of water structure. *Discuss. Faraday Soc.*, 24:133–140, 1957.

[240] Francesco Paesani and Gregory A. Voth. The properties of water: Insights from quantum simulations. *J. Phys. Chem. B*, 113(17):5702–5719, 2009.

[241] Michele Ceriotti, Wei Fang, Peter G Kusalik, Ross H McKenzie, Angelos Michaelides, Miguel A Morales, and Thomas E Markland. Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges. *Chem. Rev.*, 116(13):7529–7550, 2016.

[242] Claude Leforestier, Linda B Braly, Kun Liu, Matthew J Elrod, and Richard J Saykally. Fully coupled six-dimensional calculations of the water dimer vibration-rotation-tunneling states with a split wigner pseudo spectral approach. *J. Chem. Phys.*, 106(20):8527–8544, 1997.

[243] RS Fellers, LB Braly, RJ Saykally, and C Leforestier. Fully coupled six-dimensional calculations of the water dimer vibration-rotation-tunneling states with split wigner pseudospectral approach. ii. improvements and tests of additional potentials. *J. Chem. Phys.*, 110(13):6306–6318, 1999.

[244] Nir Goldman and RJ Saykally. Elucidating the role of many-body forces in liquid water. i. simulations of water clusters on the vrt (asp-w) potential surfaces. *J. Chem. Phys.*, 120(10):4777–4789, 2004.

[245] Nir Goldman, Claude Leforestier, and RJ Saykally. A 'first principles' potential energy surface for liquid water from vrt spectroscopy of water clusters. *Phil. Trans. R. Soc. A*, 363(1827):493–508, 2004.

[246] Marc Riera, Sandra E. Brown, and Francesco Paesani. Isomeric equilibria, nuclear quantum effects, and vibrational spectra of $m^+(h_2o)_{n=1-3}$ clusters, with m = li, na, k, rb, and cs, through many-body representations. *J. Phys. Chem. A*, 122(27):5811–5821, 2018.

[247] Kevin R Wilson, James G Tobin, AL Ankudinov, JJ Rehr, and RJ Saykally. Extended x-ray absorption fine structure from hydrogen atoms in water. *Phys. Rev. Lett.*, 85(20):4289, 2000.

[248] W Smith, TR Forester, and IT Todorov. The dl poly 2 user manual, version 2.19. *STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD Cheshire, UK*, 2008.

[249] Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, and Carol Willing. Jupyter notebooks - a publishing format for reproducible computational workflows. In F. Loizides and B. Schmidt, editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87–90. IOS Press, 2016.

[250] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.

[251] Matthias Seeger. Gaussian processes for machine learning. *Int. J. Neural Syst.*, 14(02):69–106, 2004.

[252] Sandra E Brown, Ionuţ Georgescu, and Vladimir A Mandelshtam. Self-consistent phonons revisited. ii. a general and efficient method for computing free energies and vibrational spectra of molecules and clusters. *J. Chem. Phys.*, 138(4):044317, 2013.

[253] SQLite Development Team. Sqlite.

[254] Justin M Turney, Andrew C Simmonett, Robert M Parrish, Edward G Hohenstein, Francesco A Evangelista, Justin T Fermann, Benjamin J Mintz, Lori A Burns, Jeremiah J Wilke, Micah L Abrams, et al. Psi4: an open-source ab initio electronic structure program. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2(4):556–565, 2012.

[255] Robert M Parrish, Lori A Burns, Daniel GA Smith, Andrew C Simmonett, A Eugene DePrince III, Edward G Hohenstein, Ugur Bozkaya, Alexander Yu Sokolov, Roberto Di Remigio, Ryan M Richard, et al. Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. *J. Chem. Theory Comput.*, 13(7):3185–3197, 2017.

[256] PostgreSQL Development Team. Postgresql.

[257] Maplesoft Development Team. Maple 2019 maplesoft, a division of waterloo maple inc., waterloo, ontario.

[258] Leonardo Dagum and Ramesh Menon. Openmp: An industry-standard api for shared-memory programming. *Comput. Sci. Eng.*, (1):46–55, 1998.

[259] Michele Ceriotti, Joshua More, and David E Manolopoulos. i-pi: A python interface for ab initio path integral molecular dynamics simulations. *Comput. Phys. Commun.*, 185(3):1019–1026, 2014.

[260] Daniel R Moberg, Shelby C Straight, and Francesco Paesani. Temperature dependence of the air/water interface revealed by polarization sensitive sum-frequency generation spectroscopy. *The Journal of Physical Chemistry B*, 122(15):4356–4365, 2018.